

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 8, 2004, 06:37:58 ; Search time 116.71 Seconds
(without alignments)
2596.205 Million cell updates/sec

Title: US-10-023-523-44
Perfect score: 2785
Sequence: 1 MKNQDKNGAKQSNPKSSP.....APSTASGGTGPQPTSARA 546

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues
Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Command line parameters:
-MODEL=frame+p2n model -DEV=xlh
-Q=/cgn2_1/USFCO_spool/US10023523/runat 08062004 063749 29963/app query.fasta_1.1422
-DB=Issued Patents NA -QMT=fastcap -SUFFIX=rni -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi
-LIST=50 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=50
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10023523 @CGN 1.1.98 @runat 08062004 063749 29963 -NCPU=6 -ICPU=3
-NO WMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*

1:	/cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2:	/cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3:	/cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4:	/cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5:	/cgn2_6/ptodata/2/ina/PTCUS_COMB.seq:*
6:	/cgn2_6/ptodata/2/ina/backfiles1.seq:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					Query		Match		Length		DB		ID		Description	
Result No.	Score	Match	Length	DB	ID	Score	Match	Length	DB	ID	Score	Match	Length	DB	ID	Description
1	2785	100.0	1638	4	US-09-616-289-46	2785	100.0	1638	4	US-09-616-289-46	2785	100.0	1638	4	US-09-616-289-46	Sequence 46, Appl
2	2693	96.7	4697	4	US-08-979-608A-17	2693	96.7	4697	4	US-08-979-608A-17	2693	96.7	4697	4	US-08-979-608A-17	Sequence 17, Appl
3	2693	96.7	4697	4	US-09-517-849-17	2693	96.7	4697	4	US-09-517-849-17	2693	96.7	4697	4	US-09-517-849-17	Sequence 17, Appl
4	2693	96.7	4697	4	US-09-616-289-17	2693	96.7	4697	4	US-09-616-289-17	2693	96.7	4697	4	US-09-616-289-17	Sequence 17, Appl
5	2552.5	91.7	4722	4	US-08-979-608A-14	2552.5	91.7	4722	4	US-08-979-608A-14	2552.5	91.7	4722	4	US-08-979-608A-14	Sequence 14, Appl
6	2552.5	91.7	4722	4	US-09-517-849-14	2552.5	91.7	4722	4	US-09-517-849-14	2552.5	91.7	4722	4	US-09-517-849-14	Sequence 14, Appl
7	2552.5	91.7	4722	4	US-09-616-289-14	2552.5	91.7	4722	4	US-09-616-289-14	2552.5	91.7	4722	4	US-09-616-289-14	Sequence 14, Appl
8	1597	57.3	1854	5	PTC-US94-01101-1	1597	57.3	1854	5	PTC-US94-01101-1	1597	57.3	1854	5	PTC-US94-01101-1	Sequence 1, Appl
9	1048.5	37.6	1191	2	US-08-968-751-5	1048.5	37.6	1191	2	US-08-968-751-5	1048.5	37.6	1191	2	US-08-968-751-5	Sequence 5, Appl
10	740.5	26.6	22355	4	US-09-616-289-51	740.5	26.6	22355	4	US-09-616-289-51	740.5	26.6	22355	4	US-09-616-289-51	Sequence 51, Appl
11	652.5	23.4	614	4	US-09-833-381-1333	652.5	23.4	614	4	US-09-833-381-1333	652.5	23.4	614	4	US-09-833-381-1333	Sequence 1333, Ap
12	283	10.2	9551	1	US-08-056-200-93	283	10.2	9551	1	US-08-056-200-93	283	10.2	9551	1	US-08-056-200-93	Sequence 93, Appl

Db 1563 GGGGCTCAAGAGCCACCTCCGCGAGGCC 1592
 RESULT 3
 US-09-517-849-17
 ; Sequence 17, Application US/09517849
 ; Patent No. 6605588
 ; GENERAL INFORMATION:
 ; APPLICANT: Lees, Ann M.
 ; Lees, Robert S.
 ; Law, Simon W.
 ; Ariona, Aribal A.
 ; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
 ; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
 ; TREATING ATHEROSCLEROSIS
 ; NUMBER OF SEQUENCES: 42
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson P.C.
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02110-2804
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/517,849
 ; FILING DATE: 02-Mar-2000
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: 08/979,608
 ; FILING DATE: 26-NOV-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Myers, Louis
 ; REGISTRATION NUMBER: 35,965
 ; REFERENCE/DOCKET NUMBER: 10797-003001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617/542-5070
 ; TELEFAX: 617/542-8906
 ; INFORMATION FOR SEQ ID NO: 17:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 4697 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; FEATURES:
 ; NAME/KEY: Coding Sequence
 ; LOCATION: 3...1592
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 17:
 US-09-517-849-17
 Alignment Scores:
 Pred. No.: 9,796-218 Length: 4697
 Score: 2693.00 Matches: 529
 Percent Similarity: 99.81% Conservative: 0
 Best Local Similarity: 99.81% Mismatches: 1
 Query Match: 96.70% Indels: 0
 DB: 4 Gaps: 0
 US-10-023-523-44 (1-546) x US-09-517-849-17 (1-4697)
 QY 17 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 36
 Db 3 AAAAGCAGCCAGGCAACCGGAAGCAGGACCCAGGAGGCCAGGAGGCCAGCCAG 62
 QY 37 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 56
 Db 63 CGGGCTCTCGTAGTGAAGAGCAGAGGTCCCGGACGACGAGGCTCTCGGAAGCCGAG 122
 QY 57 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 76

Db 123 GGGGCTCAAGCCAGAACCGGCTCAGTCTGGGGCCCTTCGTGATGTCCTCTGAGGAGCTGAGC 182
 QY 77 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnGlnGlnGlyProGly 96
 Db 183 CGCAACTCGAAGACATACCTAGCACAATCTGTGTGACATAAACCAGGGGGCCCCGGC 242
 QY 97 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 116
 Db 243 GAGGATGGGCACAGGGTGGCCGGCTGAACCCGGAAGATCGAGAGGTCCCGAGCTAT 302
 QY 117 ValAlaArgAsnGlyGluProGluProThrProValValAsnGlyGluLysGluProSer 136
 Db 303 GTGGCAAGGAATGGGAGCCCTGAACCACTCCAGTAGTCTATGAGAGAGGAGACCTCC 362
 QY 137 LysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHis 156
 Db 363 AAGGGGATCCAAACACAGAGAGATCCGGCAGAGTGACAGGTCCGAGACCGAGACCAT 422
 QY 157 ArgArgProGlnGlyLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMet 176
 Db 423 CGAAGGCCACAGGAGAGAAAAAGCCAGGGTTTGGGAAGGAGATCACGTTGCTGATG 482
 QY 177 GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuLysLysTyr 196
 Db 483 CAGACATTGAATCTCTAGTACCCAGAGGAGAGCTGGCTCTCTGTGCAAGAGTAT 542
 QY 197 AlaGluLeuGluGlnHisArgAsnSerGlnLysGlnMetLysLeuGlnLysLys 216
 Db 543 GCTGAATCTGTGAGGAGCACCAGGATTCACAGAGCAGATGAAGCTCTACAGAAAAAG 602
 QY 217 GlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeu 236
 Db 603 CAGAGCCAGCTGTGCAAGAGAGGACACCTCGCGGTGAGCACAGCAGAGCCCTCTG 662
 QY 237 AlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLys 256
 Db 663 GCCCGCAGCAAGCTTGAGAGCTATGCCGTGAGTGCAGCGGCACCAACCGCTCCCTCAAG 722
 QY 257 GluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHis 276
 Db 723 GAAGAAGGTGTGAGCGGGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 782
 QY 277 PheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSer 296
 Db 783 TTCAGGTGACACTGAATGACATTCAGTTCAGTGCAGATGAAACAGACCAATGAGCGCACTCC 842
 QY 297 LysLeuArgGlnGluAsnMetGluLeuAlaGluArgGluLysLysLeuLeuGlnTyr 316
 Db 843 AAGCTGCGCCAGAGAACATGAGGCTGGCTGAGAGGCTCAAGAAGCTGATTGAGCAGTAT 902
 QY 317 GluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeu 336
 Db 903 GAGCTCGCGAGGAGCATATCGACAAGTCTTCAACACAGAGGACCTCAACAGCAGAGCTG 962
 QY 337 ValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGln 356
 Db 963 GTGGATGCCAAGCTCCAGAGGCCAGGAGATGCTAAAGAGGAGCAGAGAGCGGACCCAG 1022
 QY 357 ArgGluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMet 376
 Db 1023 CGGGAGAGGATTTCTCTGAAAGGCGAGTAGAGTCCAGAGGATGTGTGAGTGATG 1082
 QY 377 LysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGlu 396
 Db 1083 AAGCAGCAAGAGAGACCCACCTGAAGCAACAGCTTGCCCTATACACAGAGAAGTTTCAGGAG 1142
 QY 397 PheGlnAsnThrLeuSerLysSerGluValPheThrThrPheLysGlnGlnMetGlu 416
 Db 1143 TTCAGAACACACTTTCCAAAGACGAGGAGGATTCCACCACATTCAACAGAGAGATGGAA 1202
 QY 417 LysMetThrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 436
 Db 1203 AAGATGACTAAGAGATCAAGAAGCTGAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1262

```
QY 437 GluSerSerAsnLysAlaLeuLeuGluMetAlaGluLysThrValArgAspLysGlu 456
Db 1263 GAGAGCAGCAACAGGCCCTGCTTGAGATGGCTGAGGAGAAACAGTCCGGGATAAGAA 1322
QY 457 LeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThr 476
Db 1323 CTGGAGGCCCTGCAGGTAAATCCACAGCGCTGGAGAGCTGTCCGGGACTGCAGACA 1382
QY 477 GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeu 496
Db 1383 GAGCGCAATGACCTGACACAGAGGGTACAGGACCTGAGTCTGGTGGCCAGGGCTCCCTC 1442
QY 497 ThrAspSerGlyProGluArgArgProGluGlyProGlyValAlaGlnAlaProSerSerPro 516
Db 1443 ACTGACAGTGGCTCTGAGAGAGCGCCAGAGGGCCCTGGGGCTCAAGCACCCAGCTCCCC 1502
QY 517 ArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThr 536
Db 1503 AGGGTCACAGAGCGCCTTCTACCCAGGAGGACCCGAGCAGACAGACATCAGGCCAGACT 1562
QY 537 GlyProGlnGluProThrSerAlaArgAla 546
Db 1563 GGGCCTCAAGAGCCACCTCCGCCAGGGCC 1592
```

RESULT 4

```
US-09-616-289-17
; Sequence 17, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Ajlona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR FILING DATE: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 4697
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(1592)
US-09-616-289-17
```

```
Alignment Scores:
Pred. No.: 9,796-218 Length: 4697
Score: 2693.00 Matches: 529
Percent Similarity: 99.81% Conservatives: 0
Best Local Similarity: 99.81% Mismatches: 1
Query Match: 96.70% Indels: 0
DB: 4 Gaps: 0
```

US-10-023-523-44 (1-546) x US-09-616-289-17 (1-4697)

```
QY 17 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 36
Db 3 AAAGCAGCCAGGACAAACCCGAGAGCCAGGAGCCAGGAGCGGCCAGCCAG 62
```

```
QY 37 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 56
Db 63 GCGGCTCTCTGAGTAAAGCAGAGAGTCCCGCAGCAGCAGGCTCTCTCGAAGCGGAG 122
QY 57 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluLeuSer 76
Db 123 GGGGCTCAAGCCAGAACCGCTCAGTCTGGGCCCCCTTCGTGTGATGCTCTGAGAGCTGAGC 182
QY 77 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGly 96
Db 183 CGCCAACTGGAAGACATACTGAGCACATCTGTGTGGACAATAACACAGGGGGCCCCCGC 242
QY 97 GluAspGlyAlaGlnGlyProAlaGluProGluAspAlaGluLysSerArgThrTyr 116
Db 243 GAGGATGGGGCAGACAGGGTGAGCCGGCTGAACCCAGAGATGAGAGAGTCCCGGACCTAT 302
QY 117 ValAlaArgAsnGlyGluProGluProThrProValValAsnGlyGluLysGluProSer 136
Db 303 GTGGCAAGGAATGGGAGCCTGMAACCACTCCAGTAGTCTATGGAGAGAGGACCTTCC 362
QY 137 LysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHis 156
Db 363 AAGGGGGATCCAAACACAGAGAGATCCGGCAGAGTACGAGGTCGGAGACCGAGACCAT 422
QY 157 ArgArgProGlnGluLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMet 176
Db 423 CGAGGCCACAGGAGAGAAAAGCCNAGGTTTGGGAAAGGAGATCACGTTGCTGATG 482
QY 177 GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaAlaLeuCysLysLysTyr 196
Db 483 CAGACATTTGAATACTCTGAGTACCCAGAGGAGAAGCTGGCTGCTCTGTGCAAGAAGTAT 542
QY 197 AlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLys 216
Db 543 GCTGAACCTGCTGGAGGAGCACCGGAATTCACAGAGCAGATGAAGCTCCTACAGAAAAG 602
QY 217 GlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeu 236
Db 603 CAGAGCCAGCTGGTGCAAGAGAGAGACCACTGCGCGGTGAGCAGACAGCAAGCGCTCTC 662
QY 237 AlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLys 256
Db 663 CCCCAGCAGCAAGCTTGAGAGCTATGCCGTGAGCTGAGCGGGCAACACCCGCTCCCTCA 722
QY 257 GluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHis 276
Db 723 GAAGAAGGTGTGCAGCGGGCCCCGGAGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGG 782
QY 277 PheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSer 296
Db 783 TTCCAGGTGACACTGAATGACATTCAGCTGCAGATGGAACAGCAACAATGAGCGCAACTCC 842
QY 297 LysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTyr 316
Db 843 AAGCTGCCCAAGAGAACATCGAGCTGGCTGAGAGGCTCAAGAAGCTGATTGAGCAGTAT 902
QY 317 GluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeu 336
Db 903 GAGCTGGCGAGGAGCATATCCGACAAAGTCTTCAACACAAAGGACCTTACACACAGCAGCTG 962
QY 337 ValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGln 356
Db 963 GTGGATGCCAAGCTCCAGCAGGCCCCAGGAGATGCTAAAGAGGAGGAGGAGGAGGAGGAG 1022
QY 357 ArgGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMet 376
Db 1023 CGGGAGAAGGATTTTCTCTCTGAAAGAGCAGTAGAGTCCCAGAGGATGTCTGAGCTGATG 1082
QY 377 LysGlnGlnGlnThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGlu 396
Db 1083 AAGCAGCAAGAGACCCACCTGAAGCAACAGCTTGCCTATACACAGAGAAAGTTTGGAGGAG 1142
QY 397 PheGlnAsnThrLeuSerLysSerGluValPheThrThrPheLysGlnGluMetGlu 416
```


838 GTCCAGCGAGCCGAGAGGAGGAGAGCGGAGGAGTGCCTGCACATTCACAGTG 897
QY ThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArg 299
Db ACCTCAACGACATTCAGTGCAGATGCGAGCGAGCAGCAGCAGCGCAACTCCAAAGTGGCG 957
QY GlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGlnTyrGluLeuArg 319
Db CAGGAGAACATGGAGTGGCGGAGCGGCTCAGAGAGCTGATGAGCAGTACAGGTGGCA 1017
QY 320 GluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeuValAspAla 339
Db 1018 GAAGAGCACATCGACAAGTCTTCAACACAAAGGATCTCAGCAGCAGCTGGTGGACGCC 1077
QY 340 LysLeuGlnGlnAlaGlnGluMetLysGluAlaGluArgHisGlnArgGluLys 359
Db 1078 AGCTCCAGAGCGCCGAGAGATGCTGAAGAGGAGCAGAGCGGCACCGCGGAGAG 1137
QY 360 AspPheLeuLysGlnAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGln 379
Db 1138 GACTTTCTCTGAAGAGCGGTGGAGTCCAGAGGATGCGAGCTGATGAAGCAACAG 1197
QY 380 GluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsn 399
Db 1198 GAGACCCACTGAAGCAGCAGCTTGCCCTATACACAGAGAAGTTTGAGAGTTCCAGAAC 1257
QY 400 ThrLeuSerLysSerGluValPheThrThrPheLysGlnGluMetGluLysMetThr 419
Db 1258 ACTTTTCCAAAGCAGCGAGGTGTTCACCACTTCAACAGCAATGGAAGAATGACA 1317
QY 420 LysLysIleLysLeuGluLysGluThrThrMetTyrArgSerArgTrpGluSerSer 439
Db 1318 AAGAAGATCAAGAAGTGGAGAGAGAGACCACTGTACCGTTCGGGTGGGAGAGCAGC 1377
QY 440 AsnLysAlaLeuLeuGluMetAlaGluGlnLysThrValArgAspLysGluLeuGlu 459
Db 1378 ACAGAGCCCTCTTGAGATGGCTGAGGAGAGAAACACTCCGGGACAAAGAGCTGAAGGC 1437
QY 460 LeuGlnValIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsn 479
Db 1438 CTGAGGTGAAATCCAGCGGTGGAGAGCTGTGCGGGCACTGCAGACAGAGCGCAAT 1497
QY 480 AspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlnGlySerLeuThrAspSer 499
Db 1498 GACCTGAACAGAGGTGCAGGACCTGAGTCCGGTGGCGAGGCGCCGCTCTCCAGACAGC 1557
QY 500 GlyProGluArgPro-----GluGlyPro 508
Db 1558 GGTCTGACGGGAGGCCAGAGCCGCCACCTCCAGAGCAGGGTGTGAGGGGCCCC 1617
QY 509 GlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyrProGlyAlaPro 528
Db 1618 GGGGCTCAAGTACCAACTCTCAAGGGGCCAGAGCGTTCCTGTGGCGAGGTGCACCC 1677
QY 529 SerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAlaArgAla 546
Db 1678 AGCAGAGGCTATCAGGCAGACAGGGCCCCAGGAGGCCACCTGCCACTGCC 1731

RESULT 6

US-09-517-849-14
; Sequence 14, Application US/09517849
; Patent No. 6605588
; GENERAL INFORMATION:

APPLICANT: Lees, Ann M.
Lees, Robert S.
Law, Simon W.

Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN

; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
; TREATING ATHEROSCLEROSIS

; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/517,849
FILING DATE: 02-Mar-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/979,608
FILING DATE: 26-NOV-1997
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Louis
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: 10797-003001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 4722 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 61...1731
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-517-849-14

Alignment Scores:

Pred. No.: 7,41e-206 Length: 4722
Score: 2552.50 Matches: 508
Percent Similarity: 93.19% Conservative: 12
Best Local Similarity: 91.04% Mismatches: 25
Query Match: 91.65% Indels: 13
DB: 4 Gaps: 3

US-10-023-523-44 (1-546) x US-09-517-849-14 (1-4722)

QY 1 MethLysAsnGlnAspLysAsnGlyAlaAlaLysGlnSerAsnProLysSerSerPro 20
Db 61 ATGAAGATCAAGACAAAAGAACGGGGCTGCCAAACAGCCCAACCCCAAAAGCAGCCG 120
QY 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
Db 121 GGACAGCCGGAAGCAGAGGAGGAGGAGCCAGGGCGCGCCGCGCGCGCGCGCGCC 180
QY 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
Db 181 CGAGAACCCAGAGT---GCCAGCAGCCAGCTCCCGGAGGCGGAGGGGGCTCAAGCC 237
QY 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
Db 238 AAAACTCTCAGCTGGGGCGCTCTGTGTGTCTCTGAGGAGCTGAGCCGCCAGTTGGAA 297
QY 81 AspileLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAla 100
Db 298 GACATCTAGTACTACTGTGTGGACAAACACAGGGGGCCCCGGGTGAGGATGGGGTC 357
QY 101 GlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsn 120
Db 358 CAGGGTGAGCCCTGAACTGAACCTGAAGATGCAGAGAAGTCTCGCCCTATGTGGCAAGAA 417
QY 121 GlyGluProGluPro---ThrProValValAsnGlyGluLysGluProSerLysGlyAsp 139
Db 418 GGGGAGCCGAGCGGGGACCCCAAGTAGTCAATGGCGAGAAGAGAGACCTCCAAAGGCAG 477

```

140 QY ProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgHisArgArgPro 159
147 Db CGGGGACCGAAGAGATCCGGACGAGGATGAGTCCGGAGACCGAGACCCCGGAGGCCA 537
160 QY GlnGluLysLysLysLysGlyLeuGlyLysGlyLysGlyLysLeuLeuMetGlnThrLeu 179
167 Db CAGGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 597
180 QY AsnThrLeuSerThrProGluGluLysLeuAlaAlaLeuCysLysLysLysLysLys 199
187 Db AACACCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 657
200 QY LeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGln 219
207 Db CTCGAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 717
220 QY LeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSer 239
227 Db CTGGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 777
240 QY LysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGly 259
247 Db AGCTCGAGAGCTGTCGGGAGCTGCAGCGGACACCGCTCGCTCAAGAGAGAGT 837
260 QY ValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnVal 279
267 Db GTGACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 897
280 QY ThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArg 299
287 Db ACCTCAACACATTCAGCTGCAGATGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 957
300 QY GlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGlnTyrGluLeuArg 319
307 Db CAGGAGAACATGAGCTGGCGGAGCTCAAGAGCTGATGAGCAGTACGAGTACGAGTGC 1017
320 QY GluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeuValAspAla 339
327 Db GAAGAGCACATCCACAAAGCTTCACACACAGAGATCTGACGAGCAGCTGTGTGACGCC 1077
340 QY LysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluLysGlnArgHisGlnArgLys 359
347 Db AAGCTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1137
360 QY AspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGln 379
367 Db GACTTTCTCTGAAGGAGCGCGTGGAGTCCAGAGATGTGAGATGTGATGAAGCAACAG 1197
380 QY GluThrHisLysLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsn 399
387 Db GAGACCCACTGAGCAGCAGCAGCTTGCCCTATACACAGAGAGATTTCAGGAGTTCAG 1257
400 QY ThrLeuSerLysSerGluValPheThrThrPheLysGlnGluMetGluLysMetThr 419
407 Db ACTCTTTCCAAAAGCAGCAGAGGTGTTCCACCATTCAAAACAGGAAATGGAAGATGACA 1317
420 QY LysLysLysLysLysLysGluLysGluThrThrMetTyrArgSerArgTyrGluSerSer 439
427 Db AAGAAGATCAAGAGCTGGGAGAAAGAGACACCATGTACCGTTCCCGTGGGAGAGCAGC 1377
440 QY AsnLysAlaLeuLeuGluMetAlaGluLysLysThrValArgAspLysGluLeuGly 459
447 Db AACAGGCCCTGTGTGATGTGCTGAGGAGAAACACATCCGGACAAAGAGTGTGAAGGC 1437
460 QY LeuGlnValLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 479
467 Db CTCAGGTGAAATCCAGCGCTGGAGAGCTGTCCGGGCACTGCAGACAGAGCCCAAT 1497
480 QY AspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlnGlySerLeuThrAspSer 499
487 Db GACCTGACAAAGAGGTGCAGGACCTGAGTGGCGGTGGCCAGGCGCCCTCTCCGACAGC 1557
500 QY GlyProGluArgArgPro-----GluGlyPro 508
1558 Db GTCTCTCAGCGGAGCCGAGAGCCCGCCACCATCCAAAGAGAGAGAGGTGTGAGAGGCC 1617
509 QY GlyValGlnAlaProSerSerProArgValThrGluAlaProCysTyrProGlyAlaPro 528
516 Db GGGGCTCAGTACCACTCTCCAAAGGCCACAGAGCTTCTCTGCGCAGGTGCACCC 1677
529 QY SerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAlaArgAla 546
536 Db AGCACAGGAGCATCAGCCACAGAGCGGCCCCAGAGGCCACCACTGCCACTGCC 1731

RESULT 7
US-09-616-289-14
; Sequence 14, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 4722
; TYPE: DNA
; ORGANISM: Oryctolagus cuniculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)...(1731)
US-09-616-289-14

Alignment Scores:
Pred. No.: 7,41e-206 Length: 4722
Score: 2552.50 Matches: 508
Percent Similarity: 93.19% Conservative: 12
Best Local Similarity: 91.04% Mismatches: 25
Query Match: 91.65% Indels: 13
DB: Gaps: 3

US-10-023-523-44 (1-546) x US-09-616-289-14 (1-4722)
QY 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerSerPro 20
Db 61 ATGAAGAATCAAGACAAAAGAACCGGGCTGCCAAACAGACCCCAACCCCAAGAGCGCCG 120
QY 21 GlyGlnProGluAlaGlyProGluGlyValGlnGlnArgProSerGlnAlaAlaProAla 40
Db 121 GGACAGCCGGAAGCAGAGAGCGGAGAGCCCGAGGGGGGGCGCGCGCGCGCGCGCGCGCC 180
QY 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
Db 181 CGAAGACCGGAAGT---GCCAGCAGCAGGCTCCCGGAGCGGCGGAGGGGGGCTCAAGCC 237
QY 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
Db 238 AAAAAGTCTGAGCTGGGGCGCTCTGTGATGTCTCTGAGAGAGTACCGCCGCGCTTGGAA 297
QY 81 AspileLeuSerThrTyrCysValAspAsnAsnGlnGlyProGlyGluAspGlyAla 100
```


Alignment Scores:

Pred. No.:	1.11e-125	Length:	1854
Score:	1597.00	Matches:	334
Percent Similarity:	97.66%	Conservative:	0
Best Local Similarity:	97.66%	Mismatches:	5
Query Match:	97.34%	Indels:	5
DB:	5	Gaps:	0

US-10-023-523-44 (1-546) x PCT-US94-01101-1 (1-1854)

208	LysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeu	227
1798	AAGCAGATGAGCTCCCTACAGAAAAAGCAGACGACGCTGGTGCAGAGAGAGACACACCTG	1739
228	ArgGlyGluHisSerLysAlaValLeuAlaAqSerLysLeuGluSerLeuCysArgGlu	247
1738	CGCGGTGAGCACAGCAGCC-GTCTTGCGCCGACGACGAGCTTGAGAGCTATGCGGTGAG	1680
248	LeuGlnArgHisAsnArgSerLeuLysGluGlyValGlnArgAlaArgGluGluGlu	267
1679	CTGCACGCGCACACCCCTCCCTCAAGAGAAAGAGTGTGCACGCGGCCCGGAGAGAGAG	1620
268	GluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGln	287
1619	GAGAAAGCGCAGGAGGTGACCTCGCACTTCCAGGTGCACAGTATGATTCACGTCGAG	1560
288	MetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGluLeuAlaGlu	307
1559	ATGGAACAGCACAAATGAGCCAACTCCAAAGCTCGCAGAGAACATGAGCTGGCTGAG	1500
308	ArgLeuLysLysLeuIleGluGlnTyrGluLeuAqGluGluHisIleAspLysValPhe	327
1499	AGGCTCAAGAAAGCTGATTGACGAGTATGAGCTGCGGAGGACATATCGACAAAGTCTTC	1440
328	LysHisLysAspLeuGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMet	347
1439	AAACACAGGACCTCAACACGACGCTGGTGGATGCCAGCTCCAGCAGGCCACGAGATG	1380
348	LeuLysGluAlaGluGluArgHisGlnArgGluLysAspPheLeuLeuLysGluAlaVal	367
1379	CTAAGAGAGCAGAAAGCGGCACACGCGGAGAGAGGATTTCTCTGAAGAGGCGAGTA	1320
368	GluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLeuLysGlnGlnLeu	387
1319	GAGTCCCGAGAGGTGTGAGCTGATGAGCAGCAGCAGAGAGACCCACCTGAAGCAACAGCTT	1260
388	AlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluVal	407
1259	GCCCTATACACAGAGAGTTTGAGGAGTTCCAGAACACATTTTCCAAAGCAGCGAGGTA	1200
408	PheThrThrPheLysGlnGluMetGluLysMetThrLysLysIleLysLysLeuGluLys	427
1199	TTCCACCACATTCAAGCAGGAGATGCAAAAGATGACTAAGAGATCAAGACCTGGAGAAA	1140
428	GluThrThrMetTyrArgSerArgTrpGluSerSerAsnLysAlaLeuLeuGluMetAla	447
1139	GAACACCAACCATGTACCGGTCCCGGTGGGAGAGCAGCAACAAAGGCCCTGCTTGATGGCT	1080
448	GluGluLysThrValArgAspLysGluLeuGluGlyLeuGlnValLysIleGlnArgLeu	467
1079	GAGGAGAAAACAGTCCGGGATAAAGAACTGGAGGACCTGCAAGTAAATAATCCACGGCTG	1020
468	GluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgValGlnAsp	487
1019	CAGAAAGCTGTGCCGGCAGCTGCAGACAGAGCGCAATGACCTGAACAGAGGTCACAGGC	960
488	LeuSerAlaGlyGlyGlnGlySerLeuThrAspSerGlyProGluArgArgProGlu-Gl	507
959	CTGAGTGCTGTGTGCCAGGGCTCCCTTACTGACATGTGGCCCTTGAGAGGACGACAGGGG	900
507	YProGlyAla-GlnAlaProSerSerProArgValThrGluAlaProCysTrpProGlyA	527


```
QY 76 rArgGlnLeuGluAspIleLeuSerThrTyrcysValaspAsnAsnGlnGlyProGlu 96
Db 171 CTCTCAATCAATGATATCTTCAACATCAA-----GGCTCAATTTGGTGGCACAAG 224
QY 96 yGluaspGlyAlaGlnGlyProAlaGluProGluaspAlaGluLysSerArgThrTy 116
Db 225 TAACAAGCATTCATTGGAAGAGGATGAAGCGAGTGAATTTATACAGAGAAACAGAAAT 284
QY 116 rValAlaArgAsnGlyGluProGluProThrProValValAsnGlyGluLysGluProSe 136
Db 285 GTGAGC-----CCAGCATCTGCAGCAAGATCAAGAGAGAAAT 326
QY 136 rLysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHi 156
Db 327 CCTGGGGA-----GAAGCTCGAAGAGATCCCTCTGATGTCAGCAAGATTC 374
QY 156 sArgArgProGlnGluLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMe 176
Db 375 AGAGTGCAACAGCAAGAAAGAACT-----TTAGGAAAGAGATTTTATTACTGAT 428
QY 176 tGlnThrLeuAsnThrLeuSerThrProGluLysLeuAlaLeuLysLysLysLys 196
Db 429 GCAAGCCCTAAACACCCCTTCAACCCAGAGAGAGCTGGCAGCTCTCTGTGAAGATA 488
QY 196 rAlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLys 216
Db 489 TCGTATCTCTCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 548
QY 216 sGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValle 236
Db 549 GAAGCCAGATGTGAAGAGAGAGAGTTCCTTGAGAGTGAACATAGCAGGCTATCTT 608
QY 236 uAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeu 256
Db 609 GGCAGAGCAAGAGTAGATCTCTTTGAGAGAACTTCAGCGTCACAAATAGACCTTAA 668
QY 256 sGluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHi 276
Db 669 GGAGGAAATATGACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 728
QY 276 sPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSe 296
Db 729 TTTCCAGATTACTTTAAATGAAATTCAGCCAGCTGGAGAGCATGATCATCCACACGC 788
QY 296 rLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLysLysLysLys 316
Db 789 CAAACTCCAGAGGAAACATTGAGCTGGGGAGAGAGCTTAAAGAGAGCTCATCGAACGTA 848
QY 316 rGluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLe 336
Db 849 CGCACTGAGGAGAGAGACATTTGATAGGTTCACACATAGAGAGACTGCAACAGAGCT 908
QY 336 uValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluGluArgHis 356
Db 909 CGTGATGCCAACTGTCAGCAACACACACTGATAAAGAGAGCTGATGTAACAAATCA 968
QY 356 nArgGluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMe 376
Db 969 GAGAGAGAGAGAGTTTATTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1028
QY 376 tLysGlnGlnThrHisLeuLysGlnGlnLeuAlaLeuTyThrGluLysPheGluGlu 396
Db 1029 GAAACAGCAAGAGAGTACAACTAAACAGCAGCTTCTCTTTATATGATAGTTTGAAGA 1088
QY 396 uPheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMet 416
Db 1089 ATTCCAGACTCATGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1148
QY 416 uLysMetThrLysLysIleLysLysLeuGluLysGlu 428
Db 1149 AAAGATGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1185
```

```
RESULT 10
US-09-616-289-51
; Sequence 51, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 2255
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-616-289-51

Alignment Scores:
Pred. No.: 1,01e-51 Length: 2255
Score: 740.50 Matches: 182
Percent Similarity: 40.56% Conservative: 7
Best Local Similarity: 39.06% Mismatches: 23
Query Match: 26.59% Indels: 255
DB: 4 Gaps: 4

US-10-023-523-44 (1-546) x US-09-616-289-51 (1-2255)
QY 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerPro 20
Db 2100 ATGAAGAACAAGCAAGCAAAAGACGGGGCTGCCAAACCAATCCAAATCCAAAGCAGGCCA 2159
QY 21 GlyClnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
Db 2160 GGACAAACCGAAGCAGAGCCCGAGGAGCCAGAGCGGCCCGCCAGCGGCGCTCTGTGCA 2219
QY 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu----- 56
Db 2220 GTAGACGACAGAGGTCCCGGAGCAGCAGCCAGGCTCTCGGAGCGGAGGCTGTGCGCAG 2279
QY 56 ----- 56
Db 2280 CTCTGCGTTGCCAGCGGCGGAGGAGAGCTGTGGGGTGGCGCTCGCTTCTTGACTTAC 2339
QY 56 ----- 56
Db 2340 AGCCGAGGCGCAGGTTGTCGGGAGGAGAGATGTAGANTGAGAGGACAGTGTGGGGGC 2399
QY 56 ----- 56
Db 2400 CGCGTCCCGCTCGCTCTGGGAGGAGTGGGAGGAGTGGCCCTCTAAGCAGAGAACAG 2459
QY 56 ----- 56
Db 2460 AGTTCTGAGAGAGAGCTCCGACGGGATTAAGTCAGGTGGGAGGAGCAACAGAGGACCCAGT 2519
QY 56 ----- 56
Db 2520 CAGGAATCCAGGTCCTCGGTTAGAAACACCTCAGCCACGAGCTAACTGCCCTTCTCTGT 2579
```

```
QY 56 ----- 56
Db 2580 TTGAGGCAATTTAGAAATGATCTGAATGCGCAAGAAATGGTTTTGTGGGGGGAAGGAGA 2639
QY 56 ----- 56
Db 2640 TGGACTAGAAGTTGCTCCGTGCCATCCCTGTGTGTGCTGATGCTTTTACATACCTTTATGATC 2699
QY 56 ----- 56
Db 2700 TAAACAATAATGTTCCGGTGGTAGTGAGAAATAGTTGTGTCATTTTACAAGTAAACAGACT 2759
QY 56 ----- 56
Db 2760 TAAAGAGTTAGGCAACGATTACTATAATTTCTTGATTTTAAAGATGTTTGGAACTCTAAA 2819
QY 56 ----- 56
Db 2820 TTCTGACAGAACTAGATTTGCTGAATGATCTCCATCTCTCTCTCTCTCTCTCTCTCTCTCT 2879
QY 56 ----- 56
Db 2880 AAAAAAGTTAGCAACATTTAACTCAAACTGATGATTTGGCTGGCGCTGAAATATCCC 2939
QY 57 -----Gly 57
Db 2940 AACCCAGTGGTATAATCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2999
QY 58 AlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGlnGluLeuSerArg 77
Db 3000 CT-CAGGCCAGAACGGCTCAGTCTGGGGCCCTTCGTGGATGTTCTCTGAGAGCTGAGCCGC 3058
QY 78 GlnLeuGluAspIleLeuSerThrTyrcysValAspAsnGlnGlyProGlyGlu 97
Db 3059 CAACTGGAGACATACTGAGCACATACTGTGTGGACAATAACAGGGGGGCCCGCGCAG 3118
QY 98 AspGlyAlaGlnGlyProAlaGluProGluAspAlaGluLeuSerArgThrTyrcysVal 117
Db 3119 GATGGGGCAGAGGTGAGCGCGCTGAACCCGAGATGCGAGAGAGTCCCGGACCTATGTG 3178
QY 118 AlaArgAsnGlyGluProGluProThrProValValAsnGlyGluGluSerProSerLys 137
Db 3179 GCAAGAAATGGGAGGCTGACCACTCCAGTAGTCAATGGAGAGAGGAGACCTCCAG 3238
QY 138 GlyAspProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAspHisArg 157
Db 3239 GGGGATCCAAACACAGAGAGATCCGGCAGAGTGGAGAGTCCGAGACCGAGACCATCGA 3298
QY 158 ArgProGlnGluLysLysAlaLysGlyLeuGlyLysGluLeuLeuMetGln 177
Db 3299 AGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3358
QY 178 ThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysTyrcysVal 197
Db 3359 -----GCTGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3400
QY 198 GluLeuLeu-----GluGluHisArgAsnSerGlnLysGlnMetLysLeuLeu 213
Db 3401 AGTCTGTGTGCGCAGGATTCAGAGGAAACGGTACTTCTCTCAGAGCAGC---AGTCACTC 3457
QY 214 GlnLysLysGlnSerGln 219
Db 3458 TAGTCTAATCAAGCCAG 3475
```

RESULT 11

```
US-09-833-381-1333
; Sequence 1333, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
```

```
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1333
; LENGTH: 614
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(614)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1333
```

Alignment Scores:

```
Align. No.: 1 52e-46 Length: 614
Score: 652.50 Matches: 144
Percent Similarity: 83.66% Conservative: 25
Best Local Similarity: 71.29% Mismatches: 25
Query Match: 23.43% Indels: 8
DB: 4 Gaps: 1
```

US-10-023-523-44 (1-546) x US-09-833-381-1333 (1-614)

```
QY 166 LysGlyLeuGlyLysGluLeuThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrPro 195
Db 9 AAAATTCTAGGAAAGAAAGTTTATTACTGATGCAAGCCCTAAACACCCCTTTCAACCCCA 68
QY 186 GluGluLysLeuAlaLeuCysLysLysTyrcysValAlaLeuLeuGluGluHisArgAsn 205
Db 69 CACCACAACTGGCAGCTCTCTGTAGAATATGCTGATCTTCTGGAGGAGCAGGAGT 128
QY 206 SerGlnLysGlnMetLysLeuLeuLysLysGlnSerGlnLeuValGlnGluLysAsp 225
Db 129 GTTCAGAAAGCAATGAAGATCTCTGCAAGAACCAAGCCAGATTTGTGAAGAAGAAAGTT 188
QY 226 HisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeu-Cy 245
Db 189 CACTTGCAGAGTAGAATAGCAATAGAGGCTATCTTGGCAGAGAGCAGCTAGATCTCTTTG 248
QY 245 sArgGluLeuGlnArgHisAsn--ArgSerLeuLysGluGlyValGlnArgAlaArg 264
Db 249 CAGAGAACTTCAGCGTCACAATTAAGACCGCTTAAGAGGAGAAATATGCAGCANGACNA 308
QY 265 GluGluGluGlu-LysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIle 284
Db 309 GAGGAAGAAGAACCGAGCTAAAGAACCACTGCACATTTCCAGATTACCTTAATGAAAT 368
QY 284 eGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetG1 304
Db 369 TCAAGCCAGCTGCGAGCAGCATGATCCACAACCGCAAACTCCGACAGGAAACATTTGA 428
QY 304 uLeuAlaGluArgLeuLysLysLeuLeuGluGlnTyrcysValGluLeuArgGluHisLeuAs 324
Db 429 GCTGGGGAGAGAGCTAAAGAGAGCTCATCGAACAGTACGCGACTGAGGAGAGAGCAGCATTTGA 488
QY 324 pLysValPheLysHisLysAspLeuGlnGlnGln-LeuValAspAla---LysLeuGln- 342
Db 489 TAAGGTGTTCAACATTAAGGAACCTGCAACAGCAGCTGTGTTGGATGCCCAAACTGCAAN 548
QY 343 GlnAlaGlnGluMetLeu-LysGluAlaGluGluArgHisGlnArgGluLys 359
Db 549 CAAACGACACAACTGATAAAAGGAAGCTGTTGTAATAAATCATCANAGAGAGAGA 600
```

RESULT 12

```
US-08-056-200-93
; Sequence 93, Application US/08056200
; Patent No. 5616500
; GENERAL INFORMATION:
; APPLICANT: Steinert, Peter M.
; APPLICANT: Lee, Seung-Chul
; APPLICANT: Kim, In-Gyu
```



```

QY 408 heThrThrPhelysGlnMetGluLysMetThrLysLysLysLysLysLysLysLysG 428
Db 4289 TGAAGCGCGAGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4348
QY 428 luThrThrMetTyrArgSerArgTTPGluSerSerAsnLysAlaLeuLeuMetAlaG 448
Db 4349 AG-----GAGAGCGCGCCAGCAGCAACTGAGGCGCGCAGCAGCAGGAAAGCGCG 4396
QY 448 luGluLysThrValArgAspLysGluLeuGluGluLeuValLysLysLysLysLysG 468
Db 4397 AGCAGCGGTGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGCGC 4452
QY 468 luLysLeuLysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgValGlnAspL 488
Db 4453 -----GAGCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4483
QY 488 euSerAlaGlyGlnGlnGlySerLeuThrAspSerGlyProGluArgArgProGluGlyP 508
Db 4484 TAGCTGA-GGAGGAGCAGGAACA-----GGCC 4509
QY 508 roGlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyrProGlyAlaP 528
Db 4510 CGGAGCGGATTAAGAGCGCGATCCGAAGTGGCAGTGCAGCTAGAAAGCGAGGCGCCGAC 4569
QY 528 roSerThrGluAlaSerGlyGlnThrGlyPro 538
Db 4570 GCAGCGCAAGCAAGTCTTACTCGAGGCGCC 4601

RESULT 13
US-08-800-644-93
; Sequence 93, Application US/08800644
; Patent No. 5958752
; GENERAL INFORMATION:
; APPLICANT: Steinert, Peter M.
; APPLICANT: Lee, Seung-Chul
; APPLICANT: Kim, In-Gyu
; APPLICANT: Chung, Soo-Il
; APPLICANT: Park, Sang-Chul
; TITLE OF INVENTION: Trichosyalin and Transglutaminase-3 and
; TITLE OF INVENTION: Methods of Using Same
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/800,644
; FILING DATE: 14-FEB-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/056,200
; FILING DATE: 30-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fredrick, Michael F.
; REGISTRATION NUMBER: 36,799
; REFERENCE/DOCKET NUMBER: NIH054.001A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (714) 760-0404
; TELEFAX: (714) 760-9502
; INFORMATION FOR SEQ ID NO: 93:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9551 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```

```

; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1507..1644
; FEATURE:
; NAME/KEY: intron
; LOCATION: 1645..2511
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2512..8070
; US-08-800-644-93

Alignment Scores:
Pred. No.: 139e-13 Length: 9551
Score: 283.00 Matches: 140
Percent Similarity: 41.43% Conservative: 97
Best Local Similarity: 24.48% Mismatches: 238
Query Match: 10.16% Indels: 99
DB: 2 Gaps: 19

US-10-023-523-44 (1-546) x US-08-800-644-93 (1-9551)

QY 10 AlaAlaLysGlnSerAsnProLysSerSerProGlyGlnProGluAlaGlyProGluGly 29
Db 3054 GCAGCAAGGCGAGAGCGCCCAAGACAGAGAGTGT-----CCAGGAGGA 3095
QY 30 AlaGlnGluArgProSerGlnAlaAlaProAlaVal-----GluAlaGlu 44
Db 3096 AGAAGAGAAAGAGTGGAGGAGGCGCGAGACAGAGTCTCCGGAAGAGAAAGAGAGTTCGA 3155
QY 45 GlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAlaArgThrAlaGln 64
Db 3156 GGAAGAGAGCGCGCAAGAGAGAGTCTCCAGGAGGAAGAGAGCAGCAGTACGGAAGCT 3215
QY 65 SerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGluAspLeuSer 84
Db 3216 GGACCGCAAGAGCTGAGGAGGAGGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3275
QY 85 ThrTyrCysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAlaGlnGly-GluPr 104
Db 3276 GCGCGAGCAGCAACTAAGCGCAAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3335
QY 104 cAlaGluProGluAsp-----AlaGluLysSerArgThrTyrValAlaAr 119
Db 3336 GCGCGAGCAGGAGGAGGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3395
QY 119 gAsnGlyGluProGluProThrProValValAsnGlyGluLysGluProSerLysGlyAs 139
Db 3396 CGAGCAGGAGGAGGAGGCGCGAGCAGCAGCTGAGCGCGCGAGCAGGAGGAGGAGGAGGAGGAG 3455
QY 139 pProAsnThrGluGluLeuArg-----GlnSerAspGluValGlyAspAspHisAr 157
Db 3456 G-----CAGCAGCTGAGGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3506
QY 157 gArgProGlnGluLysLysLysAlaLysGlyLeuGlyLysGluLeuMetGlu 177
Db 3507 GCGCGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3536
QY 177 nThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuLysLysLysLysLys 197
Db 3537 GCAGCTGAGGCGCGA-GCAGCAGCTGAGGCGCGAGCAGCAGCTGAGGCGCGAGGAGGAGGAGGAG 3595
QY 197 aGluLeuLeu-GluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysG 217
Db 3596 TGAGCGCGAGCAGCAGCTGAGGCGCGAGCAGCAGCTGAGGCGCGAGGAGGAGGAGGAGGAGGAG 3655
QY 217 lnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuA 237
Db 3656 GCGAGCAGCAGCTGAGGCGCGAGCAGCAGCTGAGGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAG 3706
QY 237 laArgSerLysLeuGluSerLysCysArgGluLeuGlnArgHisAsnArgSerLeuLysG 257

```


Db 744 CATCCAGCTCTCCGACTCGAAGAAAGAGCTGCAGGACTTTCCAGCACCGGTGGAGCTCT 803
QY 94 GlyProGluValAspGlyAlaGlnGlyGluProAlaGluPro----- 107
Db 804 GGAAGAGGGGAAGAGAGTTCACAGAGGAGATCCAGAGGAGATCGAGAACTCACCAGGAGTACGAGGA 863
QY 108 -----GluAspAlaGluLysSerArgThrTyrValAla 118
Db 864 GAAGCGCGCGCTTATGATAAATCTGAAAGAACCAAGAACAGGCTTCAGCAGGAGCTGGA 923
QY 119 ArgAsn-----GlyGluProGluProThrProValValAsnGlyGluGluLeu 134
Db 924 CGACCTGGTGTGTTGATTGTTGACCAACAGCGGCACTCGTCCAACTCGTGAAGAAAGACCA 983
QY 135 Pro-----SerLysGlyAspProAsnThrGluGluLeuArgGlnSer 148
Db 984 GAGGAATTTGATCAGTTGTTAGCCGAGGAGAAACATCTCTTCCAAATACGC----- 1037
QY 149 AspGluValGly-AspArgAspHisArgProGlnGluLysLysLysAlaLysGlyLe 168
Db 1038 -----GGATGAGAGGACAGAGCTGAGCAGAGCCAGCGAGAG----- 1077
QY 168 uGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGlu 198
Db 1078 -----GAAACCAA 1085
QY 188 sLeuAlaLeuCysLysLysTyraLaGluLeuLeuGluGluHisArgAsnSerGlnLy 208
Db 1086 GGCCTGTCTCCGTGCGGCCCTTGAAGAGCGCTTGAAGCCAAAGAGAACTCGAGCG 1145
QY 208 sGlnMetLysLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuAr 228
Db 1146 GACCAACAAATGCTCAAAAGCCGCAATGGAAGACCTGTGCTGAGTCCCAAGATGACGTGG 1205
QY 228 gGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLe 248
Db 1206 CAAGACGCTCATGAGCTGGAGAGTCCCAAGGGCGCTGGAGAGGAGAT 1265
QY 248 uGlnArgHisAsnArgSerLeuLysGluGlyValGlnArgAlaArgGluGluGlu 268
Db 1266 GAAGACGAGCTGGAAGAGCTGGAGAGCGAG-----CTGCAAGCTCGGAGGAGCGCAA 1319
QY 268 uLysArgLysGluValThr-----SerHisPheGlnValThrLeuAsnAs 283
Db 1320 ACTGCGCTGGAAGTCAACATCAGCGCTCAAGGGCCAGTTCGAA-----AGGGA 1370
QY 283 pLeGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMe 303
Db 1371 TCTCCAAGCCCGGAGGAGAGATGAGGAGAAAGAGGAGGCAACTGCAGAGACAGCTTCA 1430
QY 303 tGlu-----LeuAlaGluArgLeuLy 310
Db 1431 CGAGTATGAGCGGAAGTGGAGAGAGCGAGAAAGCAAGCTGCTGCGAGCTGCAGCAA 1490
QY 310 sLysLeuIleGlu-----GlnTyrGluLeu----- 318
Db 1491 GAAGAGCTGGAAGGGGAGCTGAAAGACTGAGAGCTTCAGCGGAGCTGCTGCATCAAGGG 1550
QY 319 -ArgGluGluHisIleAspLysValPheLys-----HisLysAspLeuGlnG 334
Db 1551 GAGGAGGAGGAGCAATCAAGCACACTGCAAGAACTGCAGGCTCAGATGAAGACTTTCAAAG 1610
QY 334 nGlnLeuValAspAlaLys----- 340
Db 1611 AGAGCTGGAAGATGCCCGGCTCCAGAGATGAGATCTTCCACAGCCAAAGAAATGA 1670
QY 341 -----LeuGlnAlaGlnGluMetLeuLysGluAl 351
Db 1671 GAAGAAAGCAAGAGCTTGGAGAGCAGACTCATGCTGAGCTCAAGAGGAGCTCGCGCGCG 1730
QY 351 aGluGluArgHisGlnArg-----GluLysAspPheLeuLeuLysGluAlaValG 368

Db 1731 TGAGGGCTCGCAAAACAAGCGGACCTCGAAGAGGAGGAACTGGCAGAGGAGCTGCCAG 1790
QY 368 uSerGlnArgMetCysGluLeuMetLysGlnGlnGlnThrHisLeuLysGlnGlnLeuAl 388
Db 1791 TAGCTCTCGCGGAAGAACGCACTCCAGAGCAGAGAACGCGCTCGAGGCGCGGATCGC 1850
QY 388 aLeuTyrThrGluLysPheGluPheGlnAsnThrLeuSerLysSerGluValPb 408
Db 1851 CCAGCTGGAGAGAGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1905
QY 408 eThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLysLysLysLysL 428
Db 1906 -----CGGTCCGCAAGCCACACAGCAGGCGGCGGAGGAGGAGGAGGAGGAGGAG 1949
QY 428 uThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeuGluMet---Al 447
Db 1950 CTGGCCACAGAGCGAGCAGCGCCGAGAGAGATGAGATGCGCGCAGCAGCTCGAGCG 2009
QY 447 aGluGluLysThrValArgAspLys-----GluLeuGluGly----- 459
Db 2010 GCAGAAACAAGAGCTCGGAGCAAGCTCCACGAGATGGAGGGCGCGCTCAAGTCCAAGTT 2069
QY 460 -----LeuGlnValLysIleGlnArgLeuGluLysLeu----- 470
Db 2070 CAAGTCCACCATCGCGCGCTGGAGGCGCAAGATTCACAGCTGGAGGAGGAGGAGGAGGAG 2129
QY 471 ----CysArgAlaLeuGlnThrGluArgAsnAspLeu-----AsnLysArgValG 486
Db 2130 GGAGGCGAGAGAGAAACAGGCGCACCAAGCTCGCTGAAGCAGAGAAACAAGAGCTGAA 2189
QY 486 nAspLeuSerAlaGlyGlyGlnGlySerLeuThrAsp-----SerG 500
Db 2190 GGAATCTT-CTCCAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2248
QY 500 yProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrG 520
Db 2249 CAGAAAGGCAATGCGCGGCTCAAGCAGCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2308
QY 520 uAlaProCysTyrProGlyAlaProSerThrGluAlaSerGly 534
Db 2309 AGTCCAGCGATCAAGCCCAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2351

RESULT 15
US-08-742-923A-5
; Sequence 5, Application US/08742923A
; Patent No. 5869611
; GENERAL INFORMATION:
; APPLICANT: Liu Pu
; APPLICANT: Collins, Francis S.
; APPLICANT: Siciliano, Michael J.
; APPLICANT: Claxton, David
; TITLE OF INVENTION: Markers for Detection of Chromosome 16
; TITLE OF INVENTION: Rearrangements
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
; STREET: P.O. Box 828
; CITY: Bloomfield Hills
; STATE: MI
; COUNTRY: USA
; ZIP: 48303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08742,923A
; FILING DATE: No. 5869611 member 1, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, DeAnn F.
; REGISTRATION NUMBER: 36683

REFERENCE/DOCKET NUMBER: 2115-00869DVC
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (810) 641-1600
 TELEFAX: (810) 641-0270
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2680 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 INDIVIDUAL ISOLATE: Sample 2
 TISSUE TYPE: Acute myelomonocytic leukemia, M4b0
 TISSUE TYPE: subtype (inv16)
 POSITION IN GENOME:
 CHROMOSOME/SEGMENT: 16(inv(16)(p13q22))
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..2451
 US-08-742-923A-5

Alignment Scores:
 Pred. No.: 4,35e-14 Length: 2680
 Score: 279.50 Matches: 146
 Percent Similarity: 38.90% Conservative: 101
 Best Local Similarity: 22.99% Mismatches: 225
 Query Match: 10.04% Indels: 164
 DB: 2 Gaps: 25

US-10-023-523-44 (1-546) x US-08-742-923A-5 (1-2680)

QY 19 SerProGlyGluPro-----GluAlaGlyProGluGlyAlaGlnGluArgPro 34
 DB 582 GTCTCAGGACACCCAGGAGTGTCTTCAAGAGAAACCCGGCAGAGCTCAA---CGTCTC 638
 QY 35 SerGlnAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLys 54
 DB 639 TACGAGCTGCCAGCT-----GGAGGAGGAGCGGACAGCTTCAAGA 683
 QY 55 ProGluGly---AlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGlu 73
 DB 684 CCAGCTGGACGAGGAGATGGAGCCCAAGCAGAACCTGGAGCGCCCATCTCCACTCTCAA 743
 QY 74 GluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGly 93
 DB 744 CATCCAGCTCTCCGACTCGAAGAGAGAGCTGAGAGACTTTGCCAGCACCGTGGAGAGCTCT 803
 QY 94 GlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluPro----- 107
 DB 804 GGAAGAGGGGGAAGAAGAGTTCCAGAAGAGATCCAGAACCTCACCCAGCAGTACGAGA 863
 QY 108 -----GluAspAlaGluLysSerArgThrTyrValala 118
 DB 864 GAAGCGCGCGTTATGATAAAGTGAAGAACCAAGAACAGAGCTTCAGCAGGAGCTGGA 923
 QY 119 ArgAsn-----GlyGluProGluProThrProValValAsnGlyGluLysGlu 134
 DB 924 CGACCTGGTGTGTGTTGTCGACCAACAGCGGCACTCGTGTCCAACTTGGAAAGAACGA 983
 QY 135 Pro-----SerLysGlyAspProAsnThrGluGluIleArgGlnSer 148
 DB 984 GAGGAAATTTGATCAGTTGTTAGCCGAGGAGAAAAACATCTCTCCAAATACGC----- 1037
 QY 149 AspGluValGlyAspArgAspHisArgArgProGlnGluLysLysAlaLysGlyLe 168
 DB 1038 -----GGATGAGAGGACAGAGCTGAGCAGAGAGCCAGGAGAG- 1077
 QY 168 uGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLys 188

DB 1078 -----GAAACCAA 1085
 QY 188 sLeuAlaLeuCysLysTyrAlaGluLeuLeuGluHisArgAsnSerGlnLys 208
 DB 1086 GGCCTCTCTCCCTGGCTGGGCGCTTGAAGAGCGCTTGAAGAGCAAGAGAACTCGAGCG 1145
 QY 208 sGlnMetLysLeuLeuGlnLysGlnSerGlnLeuValGlnGlnLysAspHisLeuAr 228
 DB 1146 GACCAACAAATGCTCAAGCCGAAATGGAGACCTGTGCTCAGCTCAGGATGACGTGGG 1205
 QY 228 sglyGluHisSerLysAlaValLeuAlaArgSerLysLeuLeuSerLysCysArgGluLe 248
 DB 1206 CAAGAACGCTCCATGAGCTGGAGAGTCCAAGCGGCGCTGAGACCCAGATGGAGGAGAT 1265
 QY 248 uGlnArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGluGluGlu 268
 DB 1266 GAAGACGAGCTGGAGAGCTGGAGGACGAG-----CTGCAAGCTTCGAGGAGCGCAA 1319
 QY 268 uLysArgLysGluValThr-----SerHisPheGlnValThrLeuAsnAs 283
 DB 1320 ACTGGCGCTGGAAGTCAACATGACGAGCGCTCAAGGCGCAGTTCGAA-----AGGA 1370
 QY 283 pIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMe 303
 DB 1371 TCTCAAGCCCGGACGAGCAGATGAGAGAGAGAGGAGGAGCACTGAGAGAGAGCTTCA 1430
 QY 303 tGlu-----LeuAlaGluArgLeuLys 310
 DB 1431 CGAGTATGAGACGGAACCTGGAAGACGAGCGAAGAACAGCTGCCCTGACGCTGCAGCAA 1490
 QY 310 sLysLeuIleGlu-----GlnTyrGluLeu----- 318
 DB 1491 GAAGAAGCTGAAGGGGACCTGAAAGACCTGGAGCTTCAGGCGGACCTCTGCCATCAAGG 1550
 QY 319 ArgGluGluHisIleAspLysValPheLys-----HisLysAspLeuGlnGlu 334
 DB 1551 GAGGAGAGAGCCATCAAGCAGCTACCGCAACTCAGGCTCAGATGAAGGACTTTCAAG 1610
 QY 334 nGlnLeuValAspAlaLys----- 340
 DB 1611 AGACTGGAAGATCCCGTGTCTCCAGATGAGATCTTTGCCACACGCAAGAGAAATGA 1670
 QY 341 -----LeuGlnGlnAlaGlnGluMetLeuLysGluAl 351
 DB 1671 GAAGAAACCAAGAGCTTGAAGCAGACCTCATGCACTACAAGAGGACCTCCGCCGCGC 1730
 QY 351 aGluGluArgHisGlnArg-----GluLysAspPheLeuLysGluAlaValG 368
 DB 1731 TGAGAGGCTCGCAACAAAGCGGACCTCGAGAGGAGAACTGGCAGAGGCTGGCCAG 1790
 QY 368 uSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLysLysGlnGlnLeuAl 388
 DB 1791 TAGCTCTCGGAAGAGAACGCACTCCAGGACGAGAGCGCGCTGGAGGCGCGATCGC 1850
 QY 388 aLeuTyrThrGluLysPheGluPheGlnAsnThrLeuSerLysSerSerGluValPh 408
 DB 1851 CCAGCTGGAGAGGAGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1905
 QY 408 eThrThrPheLysGlnGluMetGluLysMetThrLysLysIleLysLysLeuGluLysG 428
 DB 1906 -----CGGGTCCGCAAGCCACACAGCAGCGCGGAGCAGCTCAGCAACGA 1949
 QY 428 uThrThrMetTyrArgSerArgTrpGluSerSerAsnLysAlaLeuLeuGluMet---Al 447
 DB 1950 GCTGGCCACAGAGCGCAGCAGCGCCCAAGAAATGAGAGTGCCCGGAGCAGCTCGAGCG 2009
 QY 447 aGluGluLysThrValArgAspLys-----GluLeuGluGly----- 459
 DB 2010 GCAGAACAGGAGCTCCGGAGCAGCTCCACGAGATGAGAGGGGGCGCTCAAGTCCAAGT 2069
 QY 460 -----LeuGlnValLysIleGlnArgLeuGluLysLeu----- 470
 DB 2070 CAAGTCCACCATCGCGCGCTGGAGGCGCAAGATTGACAGCTGAGGAGAGGAGGAGTCA 2129

QY 471 ----CysA-galaLeuGlnThr-GluArgAsnAspLeu-----AsnLysArgValG1 486
 DB 2130 GGAGCCGAGAGACAGACAGCAGCCACCAAGTCGCTGAGACAGAAAGACAGAGCTGAA 2189
 QY 486 nAspLeuSerAlaGlyGlyGlnGlySerLeuThrAsp-----SerG1 500
 DB 2190 GGAAATCTT-GCTGCAGGTGGAGGAGCAGCGCAAGATGCCCGAGCAGTACAGAGGAGCAGG 2248
 QY 500 YProGluArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrG1 520
 DB 2249 CAGAGAAAGCAATGCCAGGTCAACAGCTCAAGAGCAGCTGGAGGAGGAGGAGGAGG 2308
 QY 520 uAlaProCysTyProGlyAlaProSerThrGluAlaSerGly 534
 DB 2309 AGTCCAGCGCATCAACGCAACCGCAGGAGGTGCGAGCGGA 2351

RESULT 16

US-08-533-306A-3
 ; Sequence 3, Application US/08533306A
 ; Patent No. 5837457
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Pu
 ; APPLICANT: Collins, Francis S.
 ; APPLICANT: Siciliano, Michael J.
 ; APPLICANT: Claxton, David
 ; TITLE OF INVENTION: Markers for Detection of Chromosome 16
 ; TITLE OF INVENTION: Rearrangements
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Harries, Dickey & Pierce, P.L.C.
 ; STREET: P.O. Box 828
 ; CITY: Bloomfield Hills
 ; STATE: MI
 ; COUNTRY: USA
 ; ZIP: 48303
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/533,306A
 ; FILING DATE: September 25, 1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Smith, Deann F.
 ; REGISTRATION NUMBER: 36683
 ; REFERENCE/DOCKET NUMBER: 2115-00869COB
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (810) 641-1600
 ; TELEFAX: (810) 641-0270
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2887 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA to mRNA
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: Homo sapiens
 ; INDIVIDUAL ISOLATE: Sample 1
 ; TISSUE TYPE: Acute myelomonocytic leukemia, M4E0
 ; TISSUE TYPE: subtype (inv16)
 ; POSITION IN GENOME:
 ; CHROMOSOME/SEGMENT: 16(inv(16)(p13q22))
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..2658
 ; US-08-533-306A-3

Alignment Scores:
 Pred. No.: 4,84e-14 Length: 2887
 Score: 279.50 Matches: 146
 Percent Similarity: 38.90% Conservative: 101
 Best Local Similarity: 22.99% Mismatches: 225
 Query Match: 10.04% Indels: 164
 DB: 2 Gaps: 25
 US-10-023-523-44 (1-546) x US-08-533-306A-3 (1-2887)
 QY 19 SerProGlyGlnPro-----GluAlaGlyProGluGlyAlaGlnGluArgPro 34
 DB 789 GCTCCAGACACCCAGAGGTGCTCAAGAAGAAACCCGCGCAGAGCTCAA---CGTGC 845
 QY 35 SerGlnAlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLys 54
 DB 846 TACGAAGCTGCCAGCT-----GGAGGAGGAGCGGAGCAACAGCCCTGCAAGA 890
 QY 55 ProGluGly--AlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGlu 73
 DB 891 CCAGCTGGACGAGAGATGGAGCGCCAGCAGACCTGGAGCGCCACATCTCCACTCTCAA 950
 QY 74 GluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrcysValAspAsnAsnGlnGly 93
 DB 951 CATCCAGCTCTCCGACTCGAAGAAGAGCTGCAGAGCTTTGCCAGCAGCACCGCTGGAAGCTCT 1010
 QY 94 GlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluPro----- 107
 DB 1011 GGAAGAGGGGAAGAAGAGGTTCAGAAAGAGAGATCGAGAACCTCACCAGCAGTACGAGGA 1070
 QY 108 -----GluAspAlaGlyLysSerArgThrTyrcysValAla 118
 DB 1071 GAAGCGCGCGCTTATGATAAACTGGAAAAGACCAAGACAGGCTTCAGCAGAGCTGGA 1130
 QY 119 ArgAsn-----GlyGluProGluProThrProValValAsnGlyGluLysGlu 134
 DB 1131 CGACCTGTTGTTGATTGATGACCAACAGCGCGCAACTCGTGTCCAACTCTGAAAAGAGCA 1190
 QY 135 Pro-----SerLysGlyAspProAsnThrGluGluIleArgGlnSer 148
 DB 1191 GAGGAATTTGATCAGTTGTTAGCCGAGGAGAAAACATCTCTTCCAAATACGC----- 1244
 QY 149 AspGluValGly-AspArgAspHisArgArgProGlnGlnLysLysLysLysLysLys 168
 DB 1245 -----GGATGAGAGGAGCAGAGCTGAGGAGCAGAGCAGGAGGAGGAGGAGG 1284
 QY 168 uGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLys 188
 DB 1285 -----GAAACCAA 1292
 QY 188 sLeuAlaAlaLeuCysLysLysTyraAlaGluLeuLeuGluGluHisArgAsnSerGlnLys 208
 DB 1293 GGCCCTGCTCCCTCGGCGCCCTTGAAGAGAGGCGCTTGAAGAGCAGGAGGAGGAGGAGG 1352
 QY 208 sGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGlnLysAspHisLeuAr 228
 DB 1353 GACCAACAATGCTCAAGCCGAAATGGAAGAGCCTGCTCAGCTCCAGATGACGTGGG 1412
 QY 228 gglyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuSerCysArgGluLe 248
 DB 1413 CAAGAACGTCATGAGCTGGAGAGTCCAAGCGCGCGCTCGAGAGCAGAGGAGGAGGAGAT 1472
 QY 248 uGlnArgHisAsnArgSerLeuLysGluGlyValGlnArgAlaArgGluGluGluGlu 268
 DB 1473 GAAGCAGCAGCTGGAAGAGCTGGAGGAGCAGG-----CTCAAGCCTTCGGAGGAGCGCAA 1526
 QY 268 uLysArgLysGluValThr-----SerHisPheGlnValThrLeuAsnAs 283
 DB 1527 ACTCGGCTGGAAGTCAACATGCAGCGCGCTCAAGGGCGCAGTTCGAA-----AGGGA 1577
 QY 283 pIleGlnLeuGlnMetGlnGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMe 303
 DB 1578 TCTCAAAGCCGGGACAGCAGAGATGAGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1637

APPLICANT: Claxton, David
TITLE OF INVENTION: Markers for Detection of Chromosome 16
TITLE OF INVENTION: Rearrangements
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Harriss, Dickey & Pierce, P.L.C.
STREET: P.O. Box 828
CITY: Bloomfield Hills
STATE: MI
COUNTRY: USA
ZIP: 48303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/742,923A
APPLICATION NUMBER: US/08/742,923A
FILING DATE: No. 5869611ember 1, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Smith, DeAnn P.
REGISTRATION NUMBER: 36683
REFERENCE/DOCKET NUMBER: 2115-00869DVC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810) 641-1600
TELEFAX: (810) 641-0270
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2887 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
INDIVIDUAL ISOLATE: Sample 1
TISSUE TYPE: Acute myelomonocytic leukemia, M4b0
TISSUE TYPE: subtype (inv16)
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 16(inv16)(p13q22)]
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2658
US-08-742-923A-3
Alignment Scores:
Pred. No.: 4,84e-14 Length: 2887
Score: 279.50 Matches: 146
Percent Similarity: 38.30% Conservative: 101
Best Local Similarity: 22.99% Mismatches: 225
Query Match: 10.04% Indels: 164
DB: 2 Gaps: 25
US-10-023-523-44 (1-546) x US-08-742-923A-3 (1-2887)
QY 19 SerProGlyGlnPro-----GluAlaGlyProGluGlyAlaGlnGluArgPro 34
Db 789 GCTCCAGGACACCCAGAGTTGCTTCAAGAGAAACCCGCGAGAGCTCAA---CGTGC 845
QY 35 SerGlnAlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLys 54
Db 846 TAGAAGCTGGCCAGCT-----GGAGGAGGAGCGGAACAGCTGCAAGA 890
QY 55 ProGluGly---AlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGlu 73
Db 891 CCAGCTGGACGAGGAGATGGAGGCCAAGCAGACCTGGAGCGCCACATCTCCACTCTCAA 950
QY 74 GluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnGlnGly 93
Db 951 CATCCAGCTCTCCGACTCGAAGAGAGCTGTCAGAGCTTTGCCAGCACCCTGGAAGCTCT 1010

RESULT 17
US-08-742-923A-3
; Sequence 3, Application US/08742923A
; Patent No. 5869611
; GENERAL INFORMATION:
; APPLICANT: Liu, Pu
; APPLICANT: Collins, Francis S.
; APPLICANT: Siciliano, Michael J.

Qy	94	GlyProGluuAspGlyAlaGlnGluProAlaGluPro	107
Db	1011	GGAAGAGGGGAAGAAGAGTTTCCAGAAAGGAGAGATCGAGAAGCTCACCAGCAGTACGAGGA	1070
Qy	108	-----GluAspAlaGluLysSerArgThrTyrValAla	118
Db	1071	GAAGCGCGCGCTTATGATAAACTGGAAGAAGACCAAGACAGCTTCAGCAGAGCTGGA	1130
Qy	119	ArgAsn-----GlyGluProGluProThrProValValAsnGlyGluLysGlu	134
Db	1131	CGACCTGGTTGTTGATTGGACACACCGCGCACTCGTCTCCAACCTGCGAAAGAAGCA	1190
Qy	135	Pro-----SerLysGlyAspProAsnThrGluGluLeuArgGlnSer	148
Db	1191	GAGGAATTTGATCAGTTGTAGCCGAGGAGAAAACATCTCTTCCAAATACGC	1244
Qy	149	AspGluValGlyAspArgAspHisArgArgProGlnGluLysLysLysAlaLysGlyLe	168
Db	1245	-----GGATGAGAGGCACAGAGCTGAGGCAGAAAGCCAGGAGAG	1284
Qy	168	uGlyLysGluLeuThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLys	188
Db	1285	-----GAAACCCAA	1292
Qy	188	sLeuAlaAlaLeuCysLysLysTyrAlaGluLeuLeuGluGluHisArgAsnSerGlnLys	208
Db	1293	GGCCCTGTCTCGCTCGGCCCTTGAAGAGGCGCTTGAAGCCAAAGAGAACTCGAGCG	1352
Qy	208	sGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArg	228
Db	1353	GACCACAAATGCTCAAAGCCGAATGGAAGACCTGCTCAGCTCCAAGATCAGCTGG	1412
Qy	228	gGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLe	248
Db	1413	CAAGAAGCTCCATGAGCTGGAGAGAGTCCAAAGCGGCCCTGGAGACCCAGATGGAGAGAT	1472
Qy	248	uGlnArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGluGluGluG	268
Db	1473	GAAGACGACAGCTGGAAGAGCTGGAGGACGAG-----CTGCAAGCCTCGAGGACGCCAA	1526
Qy	268	uLysArgLysGluValThr-----SerHisPheGlnValThrLeuAsnAs	283
Db	1527	ACTGCGCTGGAAGTCAACATGCAGCGCTCAAGGCGCAGTTGCA-----AGGGA	1577
Qy	283	pIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMe	303
Db	1578	TCCTCAAGCCCGGACGACGACAGATGAGGAGAAGAGGCGCACTCGACAGACAGCTTCA	1637
Qy	303	tGlu-----LeuAlaGluArgLeuLys	310
Db	1638	CGAGTATGAGACGGAACCTGGAAGACGAGCGAAACGAACTGCGCTGGCAGCTGCAGCAA	1697
Qy	310	sLysLeuLeuGlu-----GlnTyrGluLeu	318
Db	1698	GAAGAAGCTGGAAGGGGACCTGNAACACCTTGAGCTTCAGGCGGACTCTGCCATCAGGG	1757
Qy	319	ArgGluGluHisIleAspLysValPheLys-----HisLysAspLeuGlnGln	334
Db	1758	GAGGGAGGAAGCCATCAAGCAGCTACGCAAACTGCAGGCTCAGATGAAGGACTTTCAAAG	1817
Qy	334	nGlnLeuValAspAlaLys-----	340
Db	1818	AGAGCTGGAAGATGCCCGTCTCCAGATGAGATCTTTGGCCACAGCCAAAGAGAATGA	1877
Qy	341	-----LeuGlnGlnAlaGlnGluMetLeuLysGluAl	351
Db	1878	GAAGAAAGCCAAAGAGCTTGGAAGCAGACACCTCATGCACTACAAGAGGACCTCGCCGCCG	1937
Qy	351	aGluGluArgHisGlnArg-----GluLysAspPheLeuLeuLysGluAlaValG	368
Db	1938	TCAGAGGGCTCGCAAAACAAGCGGACCTCGAGAAGGAGGAACCTGGCAGAGGAGCTCGCCAG	1997

Qy	368	userGlnArgMetCysGlutLeuMetLysGlnGlnGluThrHisLeuLysGlnGlnLeuAl	398
Db	1998	TAGCTTCTCGGGAAAGAACACACTCCAGGACGAGAAGCGCGCTCGAGGCCCGATCGC	2057
Qy	388	aLeuTyThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSergluValPh	408
Db	2058	CCAGCTGGAGGAGGAGCTGGAGGAGGAGCAGGCACAATGTGAGGCCCATGAGCGCAC	2112
Qy	408	eThrThrPheLysGlnGluMetGluLysMetThrLysLysIleLysLeuGluLysGI	428
Db	2113	-----CGGTCGGCAAAAGCCACACAGCAGGCGCGCAGCTCAGCAACGA	2156
Qy	428	uThrThrMetTyArgSerArgTrpGluSerSerAsnLysAlaLeuLeuGluMet---	447
Db	2157	GCTGGCCACAGACGGCAGCAGCGCCCACGAAGAATGAGATGCCCCGCGCAGCAGCTCGAGCG	2216
Qy	447	aGluGluLysThrValArgAsplys-----GluLeuGluGly-	459
Db	2217	GCAGAACAGGAGCTCCGGAGCAAGCTCCACGAGATGGAGGGGGCGTGCAAGTCCAAGTT	2276
Qy	460	-----LeuGlnValLysIleGlnArgLeuGluLysLeu-	470
Db	2277	CAAGTCCACCATCGCGCGCTCGAGGCCAAGATTGCACAGCTGGAGGACGAGCTCGAGCA	2336
Qy	471	----CysArgAlaLeuGlnThrGluArgAsnAspLeu-	486
Db	2337	GGAGGCCAGAGAAACAGCGCAGCCACCACCAAGTCGTGAAGCAGAAAACAGAGCTGAA	2396
Qy	486	nAspLeuSerAlaGlyGlyGlnGlySerLeuThrAsp-----SerGI	500
Db	2397	GGAAATCTT-GCTGCAGGTGGAGGACGAGCGCAAGATGCCGCGCAGTACAAGGAGCAGG	2455
Qy	500	yProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGI	520
Db	2456	CAGAGAAAGCAATGCCAGGGTCAACAGCTCAAGAGGCAGCTGGAGGAGGCAGAGGAGG	2515
Qy	520	uaLaProCysTyProGlyAlaProSerThrGluAlaSerGly	534
Db	2516	AGTCCCGAGCGCATCAACGCCAACCGCAGGAAGCTGCAGCGGGA	2558

RESULT 18
 US-09-643-597-119 ; Sequence 119, Application US/09643597
 ; Patent No. 6426072
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Fan, Liqun
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Hosken, Nancy
 ; APPLICANT: Fanger, Gary R.
 ; APPLICANT: Li, Samuel X.
 ; APPLICANT: Wang, Aijun
 ; APPLICANT: Skeiky, Yasir A.W
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: McNeill, Patricia D.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.455C11
 ; CURRENT APPLICATION NUMBER: US/09/643,597
 ; CURRENT FILING DATE: 2000-08-21
 ; NUMBER OF SEQ ID NOS: 369
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 119
 ; LENGTH: 8948
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 US-09-643-597-119

Alignment Scores:	
Pred. No.: 3.35e-13	Length: 8948
Score: 278.00	Matches: 146
Percent Similarity: 41.87%	Conservative: 114

Best local similarity: 23.51% Mismatches: 200
Query Match: 9.98% Indels: 161
DB: 4 Gaps: 28

US-10-023-523-44 (1-546) x US-09-643-597-119 (1-8948)

```
QY 9 GlyAlaLysGlnSerAspProLys-----17
DB 4508 GGCTCAGATAGACAATCTCACCAGGAAACAGGAGCTTATCTGAAGAAATAAGAGGCT 4567
QY 18 -----SerSerProGlyGlnProGluAlaGlyProGluAlaGlyProGluAlaGlyPro 34
DB 4568 CAAGAACACTCTTAAACCCAGACACAGAGAAATCT---CAGGAGGTGGAAGAGACATCCA 4624
QY 35 SerGlnAlaAlaProAlaValGluAlaGluGly-ProGlySerSerGlnAlaProArgLy 54
DB 4625 ACA-----GCAAAAGCCACTGGCTCTCAGGTGTCTCAGAGGAA 4663
QY 54 sPro-----GluGlyAlaGlnAlaArgThrAlaGlnSer-----65
DB 4664 ACAGCAGCTGGAGGTGAGCTGACACAAGTCACTCAGATGCGAACAAGAGAGCGTTAAG 4723
QY 66 -----GlyAlaLeuArgAspValSerGluGluLeuSe 76
DB 4724 ATATAGCAATCTCTTGATGATGCTGCCAAACCATCCAGGATAAACAACAGAGATAGA 4783
QY 76 rArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnGlnGlnGlyProGly 96
DB 4784 AAGG---TTAAACAACATGATCGACAAAGAAACAATGACCGG-----4823
QY 96 yGluAspGlyAlaGlnGlyGluProAlaGluProGluAlaGluLysSerArgThr--115
DB 4824 -----AAATGCTCGAAGATGAACCGGAGATTAACAAGGGTCCA 4864
QY 116 TyrValAlaAlaArgAsnGlyGluProGluProValValAsn-----GL 131
DB 4865 GTATGACCTGCAGAAACAAACAGTAGTCGACGAGGAGACAATAAACAACCTGAAGGTTCA 4924
QY 131 yGluLysGluProSerLysGlyAspProAsnThrGluGluLeuArgGlnSerAspGluVa 151
DB 4925 GGAGCAAGAACTGACACCGCTGAGGATCGACTATGAAGAGGTTCCCGAGGAGAGACTGT 4984
QY 151 lGlyAspArgAspHisArgProGlnGluLysLysLysAlaLysGlyLeuGlyLysGly 171
DB 4985 GAAGACCAAGGATATACCGCGTTCCAGAACTCTCTGAAGAGCTGCAGCTGCAGAGCA 5044
QY 171 uIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaAl 191
DB 5045 G-----AAGGTGAAGAGAGAGCTGAATCGGCTGAAGAGGACCGGCTCAGAAGACTCCTG 5098
QY 191 aLeuCysLysLysTyrAlaGluLeuGluGluHisArgAsnSerGlnLysGlnMetLy 211
DB 5099 CAAGAGGAAGAAGCTGGAGGAAGAGCTGGAAGGCTAGAGGTCCTGAGGAGCAAGC 5158
QY 211 sLeu-----LeuGlnLysLysGlnSe 218
DB 5159 CATCAAAATCAACAACCTGACCCAGCTGGAGGAGGATCCATTGTTAAGAAGAGGAG 5218
QY 218 rGln-----LeuValGlnGluLysAsp-----HisLeuArg---GlyGluHi 231
DB 5219 TGAGATGACCTCCCGCAGCAGAGGAGGACGTGTGATGGCCACCTGAGGAGGAACACAG 5278
QY 231 sSerLysAlaValLeuAlaArg-----SerLysLeuGluSerLeuCysArgGluLeuGl 249
DB 5279 GACCCAGGAGAGCTGAGGAGGCTCTCTCTCAGCTCGAGGCGCTGAGGCGGCACTTACT 5338
QY 249 nArgHisAsnArgSerLeuLysGlu-----GluGlyValGlnArgAlaAr 264
DB 5339 CCAGGAACAGGAAGTGTCAAAACAGCTCACTTGAAGATGAGCATTTCCAGAGCGCAT 5398
QY 264 gGluGluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspI 284
DB 5399 AGAAGATAAAACAGCA-----ACGTTAAATGAAG 5428
```

```
QY 284 eGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGlnAsnMetGl 304
DB 5429 CAAATAGAAATTTGAGAGCTCGAGTCTCTCAGAGAGAACCTGCACCAAGAGCACTTGTAT 5488
QY 304 uLeuAlaGluArgLeuLysLysLeu---IleGluGlnTyrGluLeuArg-----319
DB 5489 GTTAGAAGAAGAACTCGGAACCTGAGGCTGAGTACGATACCTGAGGAGAGGACGAAG 5548
QY 320 -GluGluHisIleAspLys-----ValPh 327
DB 5549 CGAAGCGGACAGTGTATAAAATGCAACCATCTTGGAACTTAAGGAGCCAGCTGCAGATCAG 5608
QY 327 elysHisLysAspLeuGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMe 347
DB 5609 CAACAACCGGACCTCGAACTCAGCGGCTGATTATGATTATACAGAGAGAGGAGGAAA 5668
QY 347 tLeuLysGluAlaGluGluArgHisGlnArgLysAspPheLeuLeuLysGluAlaVa 367
DB 5669 TTTGACAGAGAAATGAGAATTTCCAAAAGCAGGCTTTAGAGGCATCTAATAGGATTCA 5728
QY 367 lGluSerGlnArgMetCys---GluLeuMetLysGlnGlnGluThrHisLeuLysGlnGl 386
DB 5729 GGAATCAAAAGATCACTGACTACTCAGGTGTACAGGAAGAGAGAGCTTCTGTGTAAT 5788
QY 386 nLeuAlaLeuTyrThrGluLys-----PheGluGluPheGlnAsnThrLe 401
DB 5789 CAAAGTCTTGACCAAGACAAGCAAGGCTGCAGAGCTGGAGAGTGAATCGTCTGTC 5848
QY 401 uSerLysSerSerGluValPheThrThrPhelysGlnGluMetGluLysMetThrLysLy 421
DB 5849 AAAATCAACTCTAGAGCGCAGAAACACAGGCTGAAACAGCGCTCGAGTGTGAGAAACAGCA 5908
QY 421 sIleLysLys---LeuGluLysGluThrThrMetTyrArgSerArgTyrGluSerSerAs 440
DB 5909 AATTCGAATGACTGAATCAGTGGAGACTCAATATTTCCCGCAGAGAGAGGCTATTAG 5968
QY 440 nLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLys-----455
DB 5969 GAAG-----ATAGAATCGAAGAGAGAAAGAGTGGAGAGAGAGAACACAGTCTTAGGAG 6022
QY 456 -GluLeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuCysArg-----472
DB 6023 TGAGATCGAAGAGACTCCAGCAGAGATCAAGAAATTGAAGAGAGGTGAGCGGTGAAGCT 6082
QY 473 -----AlaLeuGlnThrGluArgAsnAspLeuAsnLy 483
DB 6083 GGAGGATTCACAGGGAGACACAGTCACAGTTAGAAACAGACAGCTCCCGATATCAGAG 6142
QY 483 sArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSerGlyProGluAr 503
DB 6143 GGAGATTGATAAACTC-----AGACA 6163
QY 503 gArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGluAlaProCy 523
DB 6164 GCGCCCATATGGG-----TCCATCGAGAGACCAGCACTGAGTG 6202
QY 523 s 523
DB 6203 T 6203
```

RESULT 19

US-09-480-884A-119
; Sequence 119, Application US/09480884A
; Patent No. 6482597

GENERAL INFORMATION:

APPLICANT: Wang, Tongtong
APPLICANT: Fan, Liqun
APPLICANT: Hosken, Nancy A.
APPLICANT: Kalos, Michael D.
APPLICANT: Fager, Gary R.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

FILE REFERENCE: 210121.45506
 CURRENT APPLICATION NUMBER: US/09/460,884A
 CURRENT FILING DATE: 2001-08-27
 NUMBER OF SEQ ID NOS: 330
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 119
 LENGTH: 8948
 TYPE: DNA
 ORGANISM: Homo sapien
 US-09-480-884A-119

Alignment Scores:

Pred. No.: 3,35e-13 Length: 8948
 Score: 278.00 Matches: 146
 Percent Similarity: 41.87% Conservative: 114
 Best Local Similarity: 23.51% Mismatches: 200
 Query Match: 9.98% Indels: 161
 DB: 4 Gaps: 28

US-10-023-523-44 (1-546) x US-09-480-884A-119 (1-8948)

QY 9 GlyAlaAlaLysGlnSerAsnProLys----- 17
 DB GGCTCAGATAGACAACTCTACCCGAGAAAACAGGAGCTTATCTGAAGAAATAAAGAGGCT 4567
 QY 18 -----SerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgPro 34
 DB 4568 GAAGAACACTCTAAACCAGACACAGAGAAATCT---CAGGAGGTTGGAAGACATCCA 4624
 QY 35 SerGlnAlaAlaProAlaValGluAlaGluGly-ProGlySerSerGlnAlaProArgly 54
 DB 4625 ACA-----GCAAAGGCCACTGCTCTGAGGTCTCAGAGGAA 4663
 QY 54 sPro-----GluGlyAlaGlnAlaArgThrAlaGlnSer----- 65
 DB 4664 ACAGCAGCTGGAGTTGAGCTGAGACAGTCACTCAGATCGCAACAGAGGAGCGTAAG 4723
 QY 66 -----GlyAlaLeuArgAspValSerGluGluLeuSe 76
 DB 4724 ATATAGCAATCTTTGATGATGCTGCCAAAACCATCCAGGATAAAACAGGAGATAGA 4783
 QY 76 xArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnGlnGlyGlyProGln 96
 DB 4784 AAGG---TTAAACAACACTGATCGACAAACAATAATGACCG----- 4823
 QY 96 yGluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThr-- 115
 DB 4824 -----AAATGCTGGAAGATGAAACCGCAGATTACAAAGGTCCTCA 4864
 QY 116 -TyrValAlaArgAsnGlyGluProGluProThrProValValAsn-----G1 131
 DB 4865 GTATGACCTGCAAGAAAGCAACAGTAGTCGCGAGAGCAATPAACAACATGAAGTTCA 4924
 QY 131 yGluLysGluProSerLysGlyAspProAsnThrGluLleArgGlnSerAspGluVa 151
 DB 4925 -CGAGCAAGAACTGACACGCTCAGGATCGACTATGAAAGGGTTTCCAGGAGGAGCTGT 4984
 QY 151 lGlyAspArgAspHisArgArgProGlnGluLysLysAlaLysGlyLeuGlyLysG1 171
 DB 4985 GAAGACACAGGATATCACCGGTTCCAGAACTCTCTGAAAGAGCTGCGAGCTGCAGAGCA 5044
 QY 171 uilleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaAl 191
 DB 5045 G-----AAGGTGGAAGAGGAGCTGAATCGGCTGAAGAGGACCGGTCAGAAGACTCCTG 5098
 QY 191 leuCysLysLysTyrAlaGluLeuGluGluHisArgAsnSerGlnLysGlnMetLy 211
 DB 5099 CAAGAGGAAGAGCTGGAGGAAGAGCTGGAAGGCGATGAGAGGTCGCTGAAGGAGCAAGC 5158
 QY 211 sleu-----LeuGlnLysLysGlnSe 218
 DB 5159 CATCAAAATCACCAACTGACCCAGCAGCTGGAGCAGCGATCCATTGTTAAGAGAGGAG 5218

QY 218 xGln-----LeuValGlnGluLysAsp-----HisLeuArg---GlyGluHi 231
 DB 5219 TGAGGATGACCTCGGCGACAGAGGAGCGTCTGGATGGCCACTGAGGAAAAGCAGAG 5278
 QY 231 sSerLysAlaValLeuAlaArg-----SerLysLeuGluSerLeuCysArgGluLeuG1 249
 DB 5279 GACCAGGAGAGCTGAGGAGGCTCTCTCTGAGTCCAGGCCCTCTGAGGCGGCGAGTTACT 5338
 QY 249 nArgHisAsnArgSerLeuLysGlu-----GluGlyValGlnArgAlaAr 264
 DB 5339 CCAGAAACAGAAAGTGTCAACAAAGCTCCTTGAAGATGAGCATTTCCAGAAAGCGAT 5398
 QY 264 gGluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAsp11 284
 DB 5399 AGAAGATAAAGCAGA-----AGCTTAATGAAG 5428
 QY 284 eGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGlnAsnMetG1 304
 DB 5429 CAAAATAGAAATGAGAGGCTGCAGTCTCTCACAGAGAACTGACCAAGGAGCACTTGAT 5488
 QY 304 uLeuAlaGluArgLeuLysLysLeu---lleGluGlnTyrGluLeuArg----- 319
 DB 5489 GTTAGAAGAACTGCGGAACCTGAGCTGGAGTACGATGACTGAGGAGAGACGCAAG 5548
 QY 320 -GluGluHisIleAspLys-----ValPh 327
 DB 5549 CGAAGCGGACAGTGAATAAATCAACCATCTTTGGAACCTAAGGAGCGAGCTGCAGATCAG 5608
 QY 327 eLysHisLysAspLeuGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMe 347
 DB 5609 CAACACCGGACCTGGAACCTGAGGCTGAGGCTGATTAATGATTTACAGAGAGGAGGAAA 5668
 QY 347 tLeuLysGluAlaGluArgHisGlnArgGluLysAspPheLeuLysGluAlaVa 367
 DB 5669 TTTGAGAGCAAAATTGAGAAATTCAAAAGCAGGCTTTAGAGGCTCTAATAGGATTC 5728
 QY 367 lGluSerGlnArgMetCys---GluLeuMetLysGlnGlnGlnGlnHisLeuLysGlnG1 386
 DB 5729 GGAATCAAGAAATCAGTGTACTCAGGTGTGTACAGGAAGAGAGAGCTTCTGTGTAAT 5788
 QY 386 nLeuAlaLeuTyrThrGluLys-----PheGluGluPheGlnAsnThrLe 401
 DB 5789 CAAGCTCTGAGCAACAGCAAGCAAGCTGCGAGGCTGGAGGATGAGCTGAATCGTCG 5848
 QY 401 uSerLysSerSerGluValPheThrThrPheLysGlnGlnMetGluLysMetThrLysLy 421
 DB 5849 AAAATCAACTCTAGAGCGCAAAACAGGCTGAAACAGCGCTGAGTGTGAGAAACAGCA 5908
 QY 421 sileLysLys---LeuGluLysGluThrThrMetTyrArgSerArgTyrGluSerSerAs 440
 DB 5909 AATTCAGATGACTGANTCAGTGGAGACTCAATATCCCGCAAGGAGGAGCTATTAG 5968
 QY 440 nLysAlaLeuLeuGluMetAlaGluLysThrValArgAspLys----- 455
 DB 5969 GAAG-----ATAGAATCGAAAGAGAAAGAGTGTAGAGAGAGAGAAACAGTCTTAGGAG 6022
 QY 456 -GluLeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuCysArg----- 472
 DB 6023 TGAGATCGAAAGACTCCAAGCAGATCAAGAGATTGAAGAGAGTGTGAGCGGTAACT 6082
 QY 473 -----AlaLeuGlnThrGluArgAsnAspLeuLys 483
 DB 6083 GGAGGATCTTACCAGGAGACACAGTCAGATTAGAAACAGAACGCTCCCGATATCAGAG 6142
 QY 483 sArgValGlnAspLeuSerAlaGlyGlnGlnGlnSerLeuThrAspSerGlyProGluAr 503
 DB 6143 GGAGATTGATAAATC-----AGACA 6163
 QY 503 gArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGlnAlaProCy 523
 DB 6164 GCGCCCATATGG-----TCCCATCGAGAGACCCAGACTGAGTG 6202
 QY 523 s 523

Db 6203 T 6203

RESULT 20
 US-09-542-615A-119
 ; Sequence 119, Application US/09542615A
 ; Patent No. 6518256
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Fan, Liqun
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Hosken, Nancy A.
 ; APPLICANT: Fanger, Gary R.
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.455C8
 ; CURRENT APPLICATION NUMBER: US/09/542,615A
 ; CURRENT FILING DATE: 2000-04-14
 ; NUMBER OF SEQ ID NOS: 350
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 119
 ; LENGTH: 8948
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 US-09-542-615A-119

Alignment Scores:
 Pred. No.: 3 35e-13 Length: 8948
 Score: 278.00 Matches: 146
 Percent Similarity: 41.87% Conservative: 114
 Best Local Similarity: 23.51% Mismatches: 200
 Query Match: 9.98% Indels: 161
 DB: 28 Gaps: 28

US-10-023-523-44 (1-546) x US-09-542-615A-119 (1-8948)

QY 9 GlyAlaLysGlnSerAsnProLys----- 17
 Db 4508 GGCTCAGATAGACAATCTCACCAGGAGAAACACAGGAGCTTATCTGAAGAATAAAGAGGCT 4567

QY 18 -----SerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgPro 34
 Db 4568 GAAGAACACTCTAACCCAGACACACAGAGATCTCTCAGAGGGTGGAGAGACATCCA 4624

QY 35 SerGlnAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgly 54
 Db 4625 ACA-----GCAAAAGGCCACTGGCTCTGAGGTCTCTCAGAGGAA 4663

QY 54 SPRO-----GluGlyAlaGlnAlaArgThrAlaGlnSer----- 65
 Db 4664 ACAGCAGCTGGAGGTGAGCTGAGACAAGTCACTCAGATCGGACACAGAGAGCGCTAAG 4723

QY 66 -----GlyAlaLeuArgAspValSerGluGluLeuSe 76
 Db 4724 ATATAGCAATCTCTTGATGATCTGCCAAACCATCCAGGATAAACAAGAGATAGA 4783

QY 76 rArgGlnLeuGluAspIleLeuSerThrTyrcysValAspAsnAsnGlnGlyGlyProL 96
 Db 4784 AAGG-----TTAAACAACACTGATCGACAAGAAACAAATGACCGG----- 4823

QY 96 yGluAspGlyAlaGlnGlyProAlaGluProGluAspAlaGluLysSerArgThr-- 115
 Db 4824 -----AAATGCTCGGNAGATGAACAGCGAGATTACAAAGGGTCCA 4864

QY 116 -TyrValAlaArgAsnGlyGluProGluProThrProValValAsn-----Gln 131
 Db 4865 GTATGACCTCGAAGAACCAACAGTAGTGGCGAGGAGACAATAAACAACACTGAAGTTCA 4924

QY 131 yGluLysGluProSerLysGlyAspProAsnThrGluGluIleArgGlnSerAspGluVa 151
 Db 4925 GGAGCAAGAACTGACACGCTCTAGGATCGACTATGAAGGGTTCCACAGAGAGGACTGT 4984

QY 151 lGlyAspArgAspHisArgArgProGlnGluLysLysLysAlaLysGlyLeuGlyLysGl 171
 Db 4985 GAAGGACCAAGATATCACGGGGTTCAGAACTCTCTGAAGAGCTGCAGCTCAGAGAACA 5044

QY 171 uilleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAla 191
 Db 5045 G-----AAGGTGGAGAGAGAGCTGAATCGCTGAAGAGACCGCGTCAAGACTCTGT 5098

QY 191 aLeuCysLysLysTyraLaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLy 211
 Db 5099 CAAGAGCAAGAGCTGGAGCAAGAGCTGGAAGCAATGAGGAGTCTGCTGAAGCAAGCA 5158

QY 211 sLeu-----LeuGlnLysLysGlnSe 218
 Db 5159 CATCAAAATCACCACCTGACCCAGCAGCTGGAGAGGATCCATCTTTAAGAGAGGAG 5218

QY 218 xGln-----LeuValGlnGluLysAsp-----HisLeuArg--GlyGluHi 231
 Db 5219 TGAGATGACTCTCGGCAAGAGAGGAGCTGCTGGATGGCCACTGAGGAAAGACAGAG 5278

QY 231 sSerLysAlaValLeuAlaArg-----SerLysLeuGluSerLeuGluGluLeuGl 249
 Db 5279 GACCCAGAGAGAGCTGAGGAGGCTCTCTCTGAGTCTGAGGCCCTGAGGCGGAGTTACT 5338

QY 249 nArgHisAsnArgSerLeuLysGlu-----GluGlyValGlnArgAlaAr 264
 Db 5339 CCAGAAACAGGAAAGTGTCAACCAAGCTCACTTGAGGAATGAGCAATTCAGAAAGCGAT 5398

QY 264 gGluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIl 284
 Db 5399 AGAAGATAAAGCAGA-----AGCTTAAATGAAG 5428

QY 284 eGlnLeuGlnMetGlnGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGl 304
 Db 5429 CAAATAGAAATTCAGAGGCTGCTCTCTCAGAGAACTGACCAAGAGGACCTTGAT 5488

QY 304 uLeuAlaGluArgLeuLysLysLeu-----IleGluGlnTyrrGluLeuArg-- 319
 Db 5489 GTTAGAGAGAACTGCGGAACCTGAGGCTGAGTACGATCACTGAGGAGGAGCAAG 5548

QY 320 -GluGluHisLysLys-----Valph 327
 Db 5549 CGAAGCGGACAGTATGATAAATCAACCATCTTGGAACTAAGGAGCCAGCTGCAGATCAG 5608

QY 327 eLysHisLysAspLeuGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMe 347
 Db 5609 CAACACCGGACCTTGGACCTGAGGCTGAGGCTGATTATGATTACAGAGAGAGGAGAAA 5668

QY 347 tLeuLysGluAlaGluArgHisGlnArgGluLysAspPheLeuLeuLysGluAlaVa 367
 Db 5669 TTTGAGACAGGAAATTCAGAAATTCACAAAGCAGGCTTTAGAGGCATCTAATAGGATTCA 5728

QY 367 lGluSerGlnArgMetCys-----GluLeuMetLysGlnGlnGluThrHisLeuLysGlnGl 386
 Db 5729 GGAATCAAGAAATCAGTGTACTCAGGTGTTACAGAAAGAGAGAGGCTCTCTGGTAAAT 5788

QY 386 nLeuAlaLeuTyrrThrGluLys-----PheGluGluPheGlnAsnThrLe 401
 Db 5789 CAAGTCTCTGAGCAAGACAGGCAAGGCTCAGAGGCTGAGGATGAGTGAATCGTGC 5848

QY 401 uSerLysSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLy 421
 Db 5849 AAAATCAACTCTAGAGGCGAGAAACACAGGGTGAACAGCGCTGGAGTGTGAGAAACAGCA 5908

QY 421 stLysLysLys--LeuGluLysGluThrMetTyrrArgSerArgTrpGluSerSerAs 440
 Db 5909 AATTCAGTAATCACTGAATCAGTGAAGACTCAATATCCCGCAAGGAGGAGGCTATTAG 5968

QY 440 nLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLys----- 455
 Db 5969 GAAG-----ATAGAATCCGAAAGAGAAAGAGAGAGAGAGAGACACTCTTAGAG 6022

QY 456 -GluLeuGluGluLeuGlnValIleGlnArgLeuGluLysLeuCysArg----- 472

```

Db      6023  TGAGATCGAAAGACTCCCAAGCAGAGATCAAGAGAAATTGAAGAGAGGTGCAGGCGTAAAGCT 6082
QY      473  -----AlaLeuGlnThrGluArgAsnLeuAsnly 483
Db      6083  GGAGGATTCACACGGGAGACACAGTCACAGTTAGAAACAGAACGCTCCCGATATCAGAG 6142
QY      483  sArgValGlnAspLeuSerAlaGlyGlnGlySerLeuThrAspSerGlyProGluAr 503
Db      6143  GGAGATTATAAATCTC-----AGACA 6163
QY      503  gArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGluAlaProCy 523
Db      6164  GCGCCCATATGGG-----TCCCATCGAGAGACCCAGACTGAGTG 6202
QY      523  s 523
Db      6203  T 6203

RESULT 21
US-09-606-421B-119
; Sequence 119, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Sreiky, Iasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 119
; LENGTH: 8948
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-606-421B-119

Alignment Scores:
Pred. No.: 3,35e-13 Length: 8948
Score: 278.00 Matches: 146
Percent Similarity: 41.87% Conservative: 114
Best Local Similarity: 23.51% Mismatches: 200
Query Match: 9.98% Indels: 161
DB: 4 Gaps: 28

US-10-023-523-44 (1-546) x US-09-606-421B-119 (1-8948)
QY      9  GlyAlaAlaLysGlnSerAsnProLys----- 17
Db      4508  GGGTCAGATAGACAATCTCACCAGAGAAACAGGAGCTTATCTGAAGAAATAAAGAGGCT 4567
QY      18  -----SerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgPro 34
Db      4568  GAAGAACACTCTTAACACACACAGAGAAATCT---CAGAGGGTGAAGAGACATCCA 4624
QY      35  SerGlnAlaAlaProAlaValGluAlaGluGly-ProGlySerSerGlnAlaProArgLy 54
Db      4625  ACA-----GCMAAGGCCACTGGCTCTGAGGTGTCTCTCAGAGGAA 4663
QY      54  sPro-----GluGlyAlaGlnAlaArgThrAlaGlnSer----- 65
Db      4664  ACAGCAGCTGGAGGTTGAGCTGAGACAAGTCACTCAGATGCCAACAGAGGAGCGTAAAG 4723
QY      66  -----GlyAlaLeuArgAspValSerGluGluLeuSe 76

```

```

Db      4724  ATATAAGCAATCTCTTGATGATGTGTCACAAACCATCCAGGATAAAACAAGGAGATAGA 4783
QY      76  rArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyProGly 96
Db      4784  AAGG---TTAAACAACACTGATCGACAAAGAAACAATAATGACCGG----- 4823
QY      96  yGluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGlnLysSerArgThr-- 115
Db      4824  -----AAATCGCTCGAAGATGAACACGCGAGATTACAAAGGTTCCA 4864
QY      116  TyrValAlaAlaArgAsnGlyGluProGluProThrProValValAsn-----Gly 131
Db      4865  GTATGACCTGCGAAGAACAAACAGTAGTCGACCGAGACAATAAACAATGAGGTTCA 4924
QY      131  yGluLysGluProSerLysGlyAspProAsnThrGluGluIleArgGlnSerAspGluVa 151
Db      4925  GGACCAAGAACTGACACGCTGAGGATCGACTATGAAGGGTTTCCCGAGAGGACTGT 4984
QY      151  lGlyAspArgAspHisArgArgProGlnGlnLysLysLysAlaLysGlyLeuGlyLysGly 171
Db      4985  GAAGGACCGAGATATCACGCGGTTCAGAACTCTCTGAAAGAGCTGCAGCTGCAGAGCA 5044
QY      171  uIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluLysLysLeuAla 191
Db      5045  G-----AAGGTGAAGAGGAGCTGAATCGCTGAGAGACCGCGTCAGAAAGACTCTG 5098
QY      191  aLeuCysLysLysTyrAlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLy 211
Db      5099  CAAGAGGAAGAAGCTGGGAGGAGCTGGAAGGATGAGGAGGTCTGCTGAAGGAGCAAGC 5158
QY      211  sLeu-----LeuGlnLysLysGlnSe 218
Db      5159  CATCAAAATCACCACCTGACCCAGCAGCTGGAGCAGCATCATGTTTAAAGAGAGGAG 5218
QY      218  rGln-----LeuValGlnGlnLysAsp-----HisLeuArg---GlyGluHi 231
Db      5219  TGAGGATGACCTCGGAGCAGCAGAGGAGCTGCTGGATGGCCACTGAGGGAAGAGCAG 5278
QY      231  sSerLysAlaValLeuAlaArg-----SerLysLeuGluSerLeuCysArgGluLeuGl 249
Db      5279  GACCCAGGAAGAGCTGAGGAGGTCTCTTCTGAGGTGAGGCGCTGAGGCGGCGAGTTACT 5338
QY      249  nArgHisAsnArgSerLeuLysGlu-----GluGlyValGlnArgAlaAr 264
Db      5339  CCAGGAACAGAAAGTGTCAACAAAGCTCACTTGAGGAATGAGCATTTCCAGAGGCGAT 5398
QY      264  gGluGluGluGlyLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIl 284
Db      5399  AGAAGATAAAAGCAGA-----AGCTTAAATGAAAG 5428
QY      284  eGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnLysAsnMetGl 304
Db      5429  CAAATAGAAATTTAGAGGCTGCAGTCTCTCAAGAGAACTGACCAAGGAGGACTTGAT 5488
QY      304  uLeuAlaGluArgLeuLysLysLeu---lleGluGlnTyrGluLeuArg----- 319
Db      5489  GTTAGAAGAAGAACTCGGAACCTGAGGCTGGAGTACGATGACCTGAGGAGGAGCAAG 5548
QY      320  -GluGluHisIleAspLys-----ValPh 327
Db      5549  CGAAGCGGACAGCTGATAAAATGCAACCATCTTTGGAACCTAAGGAGCCAGCTGCAGATCAG 5608
QY      327  eLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnLysMe 347
Db      5609  CAACAACCGGACCTGGAAGCTGAGGGGCTGATTAATGATTTACAGAGAGAGGAGGAAA 5668
QY      347  tLeuLysGluAlaGluArgHisGlnArgGluLysAspPheLeuLeuLysGluAlaVa 367
Db      5669  TTTGACAGCAAAATTCAGAAATTCCAAAGCGAGGCTTTAGAGGCTATTAATAGGATTC 5728
QY      367  lGluSerGlnArgMetCys---GluLeuMetLysGlnGlnGluThrHisLeuLysGlnGl 386

```

```

Db      5729 GGAATCAAGAAATCAGTGTACTCAGGTGGTACAGGAAGAGAGAGAGCCCTTCTGGTGAAT 5788
Qy      386 nLeuAlaLeuTyThrGluLys-----PheGluGluPheClnAsnThrLe 401
Db      5789 CAAAGTCTCGAGCAAGCAAGCAAGCTGCAGAGCGCTGGAGATGAGTGAATCGTGC 5848
Qy      401 uSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysly 421
Db      5849 AAAATCAACTCTAGAGCGAGAAACCCAGGTGAAACAGCCCTGGAGTGTGAGAAACAGCA 5908
Qy      421 sileLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 440
Db      5909 AATTGAGATGACCTGAATCAGTGGAGAGCTCAATATTCGCGCAGAGGAGGCTATTAG 5968
Qy      440 nLysAlaLeuLeuGluMetAlaGluGluLysLysLysLysLysLysLysLysLysLys 455
Db      5969 GAAG-----ATAGAAATCGGAAGAGAAAGAGTGAAGAGAGAGAGAAACAGTCTTAGGAG 6022
Qy      456 -GluLeuGluGluLeuGlnValLysLysLysLysLysLysLysLysLysLysLysLys 472
Db      6023 TGAGATCGAAAGACTCCAGCAGAGATCAGAGAAATGAGAGAGGTGCAGCGCTAAGCT 6082
Qy      473 -----AlaLeuGlnThrGluArgAsnAspLeuAsnly 483
Db      6083 GGAGGATTCACAGGAGAGACAGTCAACAGTTAGAAACAGACGCTCCGATATCAGAG 6142
Qy      483 sArgValGlnAspLeuSerAlaGlyGlnGlySerLeuThrAspSerGlyProGluAr 503
Db      6143 GGAGATTGATAAATC-----AGACA 6163
Qy      503 sArgProGluGluProGlyAlaGlnAlaProSerSerProArgValThrGluAlaProCy 523
Db      6164 GCGCCCATATGG-----TCCCATCGAGAGACCCAGACTGAGTG 6202

Qy      523 s 523
Db      6203 T 6203

RESULT 22
US-09-221-107-119
; Sequence 119, Application US/09221107
; Patent No. 6660838
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
; FILE REFERENCE: 210121.455C2
; CURRENT APPLICATION NUMBER: US/09/221,107
; CURRENT FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 161
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 119
; LENGTH: 8948
; TYPE: DNA
; ORGANISM: Human
US-09-221-107-119

Alignment Scores:
Pred. No.:      3,35e-13      Length:      8948
Score:          278.00      Matches:      146
Percent Similarity: 41.87%      Conservative: 114
Best Local Similarity: 23.51%      Mismatches: 200
Query Match:      9.98%      Indels:      161
DB:              4          Gaps:      28

US-10-023-523-44 (1-546) x US-09-221-107-119 (1-8948)
Qy      9 GlyAlaAlaLysGlnSerAsnProLys----- 17
Db      4508 GGTCAGATAGACAATCTCCCGAGAAACACAGAGCTTATCTGAGAAATAAAGAGGCT 4567
Qy      18 -----SerSerProGlyGlnProGluAlaGlyProGluGluAlaGlnGluArgPro 34
Db      4568 GAAGAACACTCTAAACCCAGACACAGAGAAATCT---CAGGAGGCTGGAAGAGACATCCA 4624

```

```

Qy      35 SerGlnAlaAlaProAlaValGluAlaGluGly-ProGlySerSerGlnAlaProArgLy 54
Db      4625 ACA-----GCAAAGGCCACTGGCTCTCAGGTGCTCTCAGAGAA 4663
Qy      54 spro-----GluGlyAlaGlnAlaArgThrAlaGlnSer----- 65
Db      4664 ACAGCAGCTGGAGGTGGAGTGCAGACAAGTCACTCAGATGCGACAGAGAGAGCGTAAG 4723
Qy      66 -----GlyAlaLeuArgAspValSerGluGluLeuSe 76
Db      4724 ATATAAGCAATCTCTTGTATGATGCTGCCAAAACCATCCAGGATAAAACAAGAGGATAGA 4783
Qy      76 rArgGlnLeuGluAspIleLeuSerThrTyrcysValAspAsnAsnGlnGlyGlyProGlu 96
Db      4784 AAGG-----TAAACACACTGATCGACAAAGAAACAAATGACCG----- 4823
Qy      96 yGluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThr-- 115
Db      4824 -----AAATGCTGTAAGATGMAAACCGGAGATTACAAAGGGTCCA 4864
Qy      116 -TyrValAlaAlaArgAsnGlyGluProGluProThrProValValAsn-----G1 131
Db      4865 GTATGACTGCAGAAAGCAAAACAGTAGTGCACGCGAGACAATAACAACTGAAGGTTCA 4924
Qy      131 yGluLysGluProSerLysGlyAspProAsnThrGluGluLeuArgGlnSerAspGluVa 151
Db      4925 GGAGCAAGAACTGCACACGCTGAGATCGACTATGAAAGGCTTCCAGAGAGAGACTGT 4984
Qy      151 lGlyAspArgAspHisArgArgProGlnGluLysLysLysLysLysGlyLeuGlyLysG1 171
Db      4985 GAAAGCACCAGGATATCACGCGTTCCAGAACTCTCTGAAAGAGTGCAGCTGCAGAAACA 5044
Qy      171 uIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaAl 191
Db      5045 G-----AAGGTGGAAGAGGAGCTGAATCGGCTGAAGAGGACCGCTCAGAAAGACTCTG 5098
Qy      191 aLeuCysLysLysTyraAlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLy 211
Db      5099 CAAGAGGAAGAGTGGAGGAAGAGCTGGAAGGATGAGAGGTCGCTGTAAGAGCAAGC 5158
Qy      211 sLeu-----LeuGlnLysLysGlnSe 218
Db      5159 CATCAAAATCAACCACTGACCCAGCAGCTGGAGAGCATCCATTGTTAAGAGAGAGAG 5218
Qy      218 rGln-----LeuValGlnGluLysAsp-----HisLeuArg---GlyGluHi 231
Db      5219 TGAGGATCACCTCCGGCAGCAGAGGAGCGTGTGATGGCCACCTGAGGAGAAAGACAG 5278
Qy      231 sSerLysAlaValLeuAlaArg-----SerLysLeuGluSerLeuCysArgGluLeuG1 249
Db      5279 CACCCAGGAAGAGCTGAGAGGCTCTCTTCTGAGTCCGAGGCCCTGAGGCGCGAGTTACT 5338
Qy      249 rArgHisAsnArgSerLeuLysGlu-----GluGlyValGlnArgAlaAr 264
Db      5339 CAGGAACAGGAAAGTGTCAACCAAGCTCATTGAGGAATGAGCATTTCCAGAGGCGAT 5398
Qy      264 gGluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspI1 284
Db      5399 AGAGATAAAGACAGA-----AGCTTAAATGAAAG 5428
Qy      284 eGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetG1 304
Db      5429 CAAAATAGAAATTGAGAGGCTGCTCTCTCAGAGAACCTGACCAAGAGGACACTTGTAT 5488
Qy      304 uLeuAlaGluArgLeuLysLysLeu---IleGluGlnItyrGluLeuArg----- 319
Db      5489 GTTAGAAGAGAACTGCGGAACCTGAGGCTGGATGATGATGATGATGATGATGATGATGAT 5548
Qy      320 -GluGluHisIleAspLys-----Valph 327
Db      5549 CGAAGCGGACAGTGTATNAATAATGCAACCATCTTGGAACATAAGAGCCAGCTGCAGATCAG 5608

```

327	QY	eYsHisLysAspLeuGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnLume	347
5609	Db	CAACAACCCGACCTCGAATCGCAGGGCGCTGATTAATGATTTTACAGAGAGAGGAAAA	5668
347	QY	tLeuLysGluAlaGluGluArgHisGlnArgGluLysAspPheLeuLeuLysGluAlaVa	367
5669	Db	TTTGAGACAGCAAAATGAGAAAATCCAAAGACAGGCTTTAGAGGCATCTAATAGGATTCA	5728
367	QY	lGluSerGlnArgMetCys---GluLeuMetLysGlnGlnGluThrHisLeuLysGlnGl	386
5729	Db	CGAATCAAGAATCATGCTACTCAGGTGGTACAGGAAGAGAGAGAGCGCTTCTGCTGAAAT	5788
386	QY	nLeuAlaLeuTyThrCluLys-----PheGluGluPheGlnAsnThrLe	401
5789	Db	CAAAAGTCTCGACGACAGACAGGCAAGCGCTGCAGAGGCTCGAGATGAGCTGAATCGTGC	5848
401	QY	uSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLy	421
5849	Db	AAAATCAACTCTAGAGGCAGAAAACCCAGGGTGAAACACAGCGCTGGAGTGTGAGAAAACAGCA	5908
421	QY	sIleLysLys--LeuCluLysGluThrThrMetTyArgSerArgTrpGluSerSerAs	440
5909	Db	AATTCAGATGACCTGAATCAGTGGAAAGACTCAATATCCCGCAAGAGAGGCGCTATTAG	5968
440	QY	nLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLys-----	455
5969	Db	GAAG-----ATAGAATCGGAAAAGAGAAAAGAGTGGAGAGAGAGAAACAGTCTTAGGAG	6022
456	QY	GluLeuGluGlyLeuGlnValIysIleGlnArgLeuGluLysLeuCysArg-----	472
6023	Db	TGAGATCGAAAGACTCCACACAGAGATCAAGAGAAATGGAAGAGGTGCGAGCGTAAGCT	6082
473	QY	-----AlaLeuGlnThrGluArgAsnAspLeuAsnLy	483
6083	Db	GGAGGATCTACCGGGAGACACAGTCACAGTTAGAAAACAGACGCTCCCGATATCAGAG	6142
483	QY	sArgValGlnAspLysSerAlaGlyGlyGlnGlySerLeuThrAspSerProGluAr	503
6143	Db	GGAGATTGATAAACTC-----AGACA	6163
503	QY	sArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGluAlaProCy	523
6164	Db	GCSCCCATATGGG-----TCCCATCGAGAGACCCAGCTAGTG	6202
523	QY	s	523
6203	Db	T	6203

RESULT 23

US-09-023-655-548/C
; Sequence 548, Application US/09023655
; Patent No. 6607879

; PATENT NO. 6607879
; GENERAL INFORMATION:

APPLICANT: Cocks, Benjamin G.

APPLICANT: Susan G. Stuart

APPLICANT: Jeffrey J. Seilhamer

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

INVENTOR	DATE OF INVENTION	TITLE OF INVENTION	EXPERIMENTAL DATA
...

; NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:

; CORRESPONDENCE ADDRESS:
ADDRESSEE: INOVTE PHARMACEUTICALS INC

ADDRESS: INCYTE PHARMACI
STREET. 3174 PORTER DRIVE

STREET: 3174 POR
CITY: PAJO ALTO

CITY: PALO ALTO
STATE: CALIFORNIA

STATE: CALIFORNIA
COUNTRY: USACOUNTRY: US
ZIP: 94304

ZIP: 94304
COMPUTER READABLE FORM:

```

;
; COMPUTER READABLE FORM.
;
; MEDIUM TYPE: Floppy disk

```

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

```
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
```

; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/023,655

```

, FILING DATE: HERewith
,
, CLASSIFICATION:
,
, PRIOR APPLICATION DATA:
, APPLICATION NUMBER:
, FILING DATE:
,
, CLASSIFICATION:
, ATTORNEY/AGENT INFORMATION:
, NAME: Zeller, Karen J.
, REGISTRATION NUMBER: 37,
, REFERENCE/DOCKET NUMBER:
, TELECOMMUNICATION INFORMATION:
, TELEPHONE: (650) 855-0555
, TELEFAX: (650) 845-4166
, INFORMATION FOR SEQ ID NO: 5
, SEQUENCE CHARACTERISTICS:
, LENGTH: 540 base pairs
, TYPE: nucleic acid
, STRANDEDNESS: single
, TOPOLOGY: linear
,
, IMMEDIATE SOURCE:
, LIBRARY: PANCNOT01
, CLONE: 224798
,
, US-09-023-655-548

```

Alignment Scores:	
Pred. No.:	1,518-14
Score:	273.00
Percent Similarity:	59.15%
Best Local Similarity:	44.51%
Query Match:	9.80%
DB:	4
Length:	540
Matches:	73
Conservative:	24
Mismatches:	42
Indels:	27
Gaps:	5

US-10-023-523-44 (1-546) X US-09-023-655-548 (1-540)

255	LeuLysGluGluGlyValGlnArgGluGluGluGluLysArgLysGluValThr	274
Qy		
Db		
425	CTNAAATANGANAATTTGGAACGGCAGANGTACANCTAA	381
Qy		
Db		
275	SerHisPheGln-----ValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHis	291
Qy		
Db		
380	GGCAGCTTCNCNTTATATGATGAATGTTGAA-GAATTCAGACTACCATGGCAAAAGC	322
Qy		
Db		
292	AsnGluArgAsnSerLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLys	311
Qy		
Db		
321	AATGAACCTGTTTACAACTTCAGACAGGAA---ATGGAA-----AAGATG	280
Qy		
Db		
312	LeuIleGluGlnTyrGluLeuArgGluGluHisIleAspLysValPheLysHisLysAsp	331
Qy		
Db		
279	AGCGGCNCNCGTCCGCAATT-CGCGNCAG-----	251
Qy		
Db		
332	LeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAla	351
Qy		
Db		
250	---CCACAGCAGCTCGTGGATGCCAACTGCAGCAANCGACAACCTGATTAACAGAA	194
Qy		
Db		
352	GluGluArgHisGlnArgGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArg	373
Qy		
Db		
193	GATGAAACATCATGAGACAGAGAGATTTTNTTAAAGACGACAGAAATCGAGGCAC	134
Qy		
Db		
372	MetCysGluLeuMetLysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrThr	391
Qy		
Db		
133	AAATACGNACAAATGAAACACAAACGACGAGTACAACATAAAACAGCAGCTTCTCTTATATG	74
Qy		
Db		
392	GluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluValPheThrPhe	411
Qy		
Db		
73	GATAGTTTGAGAAATCCAGACTACCATGGCNAAGCAATGAACCTGTTTACACCTTC	14
Qy		
Db		
412	LysGlnGluMet	415
Qy		
Db		
13	AGACAGGAAATG	2
Qy		
Db		

RESULT 24
US-08-938-105-2
; Sequence 2, Application US/08938105

Patent No. 6353151
 GENERAL INFORMATION:
 APPLICANT: Weinwand, Leslie A.
 APPLICANT: Vikstrom, Karen L.
 TITLE OF INVENTION: TRANSGENIC MODEL FOR HEART FAILURE
 NUMBER OF SEQUENCES: 3
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sheridan Ross P.C.
 STREET: 1700 Lincoln St., Suite 3500
 CITY: Denver
 STATE: CO
 COUNTRY: U.S.A.
 ZIP: 80203
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/938,105
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Crook, Wanneil M.
 REGISTRATION NUMBER: 31,071
 REFERENCE/DOCKET NUMBER: 3595-4
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (303) 863-9700
 TELEFAX: (303) 863-0223
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5661 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..5661
 US-08-938-105-2

Alignment Scores:
 Pred. No.: 9,018-13 Length: 5661
 Score: 269.50 Matches: 173
 Percent Similarity: 36.22% Conservative: 82
 Best Local Similarity: 24.57% Mismatches: 214
 Query Match: 9.68% Indels: 236
 DB: 4 Gaps: 31

US-10-023-523-44 (1-546) x US-08-938-105-2 (1-5661)

QY 19 SerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAla 38
 DB 2946 TCTGGAGGGATCCCTGGAGCAGGAGAGGCT----- 2978
 QY 39 ProAlaValGluAlaGluGlyProGlySerSerGlnAla----- 51
 DB 2979 -----GCCGATGACCTGGAGCGCAAGCGGAAAGCTGGAGGGTGACCTGAA 3026
 QY 52 -----ProArgLysProGluGlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArg 69
 DB 3027 GCTGACCCAGGAGAGCATCATGACCTGGAGAACGACAGCTTCAGCTGGAGGA----- 3080
 QY 70 AspValSerGluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysValAsp 89
 DB 3081 -----AAAGCTCAAGAAAGAAAGAGTTTGACATCAGTCAGCAGCAAGACAG 3122
 QY 90 AsnAsnGlnGly-----GlyProGlyGluAspGlyAla-----GlnGlyGluPro 104
 DB 3123 TAAATAGAGGACGAGCAGCGCCCTGGCCCTTCAGCTGCGAGAGAACTGAGGAAACCA 3182
 QY 105 AlaGluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsnGlyGluProGlu 124

DB 3183 GGCACGCATCGAGGAGCTGGAGGAGGA-----GCTAGAGCGGAGCGCACAGC 3230
 QY 125 ProThrProValValAsnGlyGluLysGluProSerLysGlyAspProAsnThr-GluGl 144
 DB 3231 CCGGCG-----CAAGGTGGAGAAAGTGCGCTCAGACCTGACCCGGAGCTGGAGG 3281
 QY 144 uileArgGlnSer---AspGluValGly-----AspargAspHi 156
 DB 3282 GATCAGTGAGAGGCTAGAGGAAGCCGTTGGGGCCACATCTGTCCAGATAGAGATGAACA 3341
 QY 156 sArgArgProGlnGluLysLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMe 176
 DB 3342 GAAGCGCGAGCGCGAGTTCCAGAAGATGCGGGGAGCCCTGGAGGAAGCCACGCTG----- 3396
 QY 176 tGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuLysLysLysLys 196
 DB 3397 -----CAGCATGAGGCCACAGCGCGCCCTGGCAGAGGA 3434
 QY 196 r-----AlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLys 212
 DB 3435 CGCAGACAGCGCTGGCGGAGCTGGCGGAGCAGATAGACAATCTACAGCGGGTGAAGCAGAA 3494
 QY 212 uLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHis 232
 DB 3495 GCTGGAGAAAGAGAGAGCGAGTTCAAACTGGAGCTGGATGACGTCACTCTCAGATGGA 3554
 QY 232 rLysAlaValLeuAlaArgSerLysLysLeuGluSerLeuCysArg----- 246
 DB 3555 GCAGATCATCAAGGCCCAAGCAACCTGGAGAAAGTGTCCCGGACACTGGAGGACCAGGC 3614
 QY 247 -----GluLeuGlnArgHisAsnArgSerLeuLysGluGluGlyValGl 261
 DB 3615 GAATGAATACCGGGTGAAGTTGGAAGAGCCAGCCGCTCCCTCAATGACTTCCACACACA 3674
 QY 261 nArgAlaArg----- 264
 DB 3675 GCGAGCCCAAGCTGCAGACAGAGATGGCAGTTGGCTAGGCACTGGAAGAAAGAGGC 3734
 QY 264 ----- 284
 DB 3735 ACTGATTGGCAGCTGACCGGGGCAAGCTCTCTATATCCACAGATGGAGGACCTCAA 3794
 QY 265 -----GluGluGluGluLysArgLysGluValThrSerHis---PheGlnValTh 280
 DB 3795 GAGCAGCTGGAGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3854
 QY 280 rLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArg----- 294
 DB 3855 CCGGCATGACTGCGACCTGCTGGCGGAAACAGTACGAAGAAAGAAATGGAGGCCAAGCCGA 3914
 QY 295 -----AsnSerLysLeuArg----- 299
 DB 3915 GCTGACGCTGCTGTCCAAAGCCAACTCAGAGGTGGCCAGCTGGAGGAGCCAGATATGA 3974
 QY 300 -----GlnGluAsnMetGluLeuAlaGluArgLeuLysLysLysLysLysLysLys 315
 DB 3975 GACGAGCCCATACAGAGGAGGAGGAGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4034
 QY 315 nTyrGluLeuArgGluGluHisLysLysVal----- 326
 DB 4035 GCTTCAGATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4094
 QY 327 -PheLysHisLys-----AspLeuGlnGlnLeuValAspAlaLysLeuGlnGl 343
 DB 4095 GACCAACACAGGCTGCGAGAACGAGATCGAGACCTGATGTGTGTGTGTGTGTGTGTGTGT 4154
 QY 343 nAlaGlnGluMetLeuLysGluAlaGluGluArgHisGlnArgGluLysAspPheLeuLe 363
 DB 4155 TGCG-----GCCCGCGAGCCCTGGACAAAAAGCAGAGGAACTTCACAGATCCT 4205
 QY 363 uLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlnThrHisLe 383
 DB 4206 GGCTGAG-----TGAAGCAGAGAAGTATGAGGAGTCCCATGAGCTGAGGCTGCTCCCA 4259

```

QY 383 ulysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLys 403
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4260 GAAGGAGCGCGCTCCCTGACACAGAGCTCTC---AAGCTCAAGATGCCCTATGAGGA 4316
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 403 sSerSerGluValPheThrThrPhelysGlnGluMetGluLysMetThrLysLysIle-- 422
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4317 GTCTCTGGAGCACCTCGAGACCTTCAAGCGGGGAGAACAGAACCTCCAGGAGGATCTC 4376
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 422 ----- 422
Db 4377 AGACCTGACTGACACAGCTGGGAGAGGGGTAAATGTGCACAGCTGGAGAGATCGG 4436
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 423 -LysLysLeuGluLysGluThrThrMetTyrArgSerArgTrpGluSerSerAsnLysAl 442
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4437 CAACACAGCTGGAGGTGGAGAGCTGGAAGCTGCACTGAGCTCGAGGAGGCT---GAGGC 4493
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 442 aLeuLeuGluMetAlaGluGluLysThrValArgAspLysGluLeuGluGluVal 462
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4494 CTCCCTGGAGCATGAGGAGGCGAAGATCTCCGA---GCCACGCTGGAGTTCAACCAT 4550
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 462 lLysIleGlnArgLeuGluLysLeuLysArg----- 472
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4551 CAAGGCAGAGATCGAAAGGAGAGCTGGCAGAGAGGACGAGGAGAGATGGAGCGCAACGC 4610
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 473 -----AlaLeuGlnThr-----GluAr 478
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4611 CAACACCTCGCGGTGTGTGACTCCCTACAGACCTCCCTGGATGCCGAGACACGACGCG 4670
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 478 gAsn-----AspLeuAsnLysArgValGlnAs 487
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4671 CAACGAGGCCCTCGCGGTGAAGAGAGATGGAGGCGGACCTCAATGAGATGGATCCA 4730
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 487 pLeuSerAlaGlyGlnGlnGlnThrThrAspSerGlyProGluLysArgProGluG 507
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4731 GCTCAGTCA-GGCCAATAGATAGCTC-----AGAGGCCAGAACACTTGAAGACG 4783
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 507 yPro-----GlyValaGlnAlaProSerSerProArgValThrGluAlaProCy 523
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4784 CCAAGCCCACTTGAAGGACACCCAGCTCCAGTGGATGACCC-----AGTCCGTG 4834
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 523 s-----TyrProGlyAlaProSerThrGluAlaSerGl 534
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4835 CCAATGACGACCTGAAGGAGAACATCGCCATCGTGAGCGGCGCAACACCTCTGTGCAGG 4894
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 534 yGlnThrGly 537
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4895 CGGAGCTGGA 4904
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 25

```

US-08-875-435B-5
; Sequence 5, Application US/08875435B
; Patent No. 6593304
; GENERAL INFORMATION:
; APPLICANT: Hasegawa, Kazuhide
; APPLICANT: Arakawa, Emi
; APPLICANT: Oda, Shoji
; APPLICANT: Matsuda, Yuzuru
; APPLICANT: Takahashi, Katsuhito
; APPLICANT: Sugahara, Michihiro
; APPLICANT: Ishiyama, Haruo
; TITLE OF INVENTION: RECOMBINANT DNA COMPRISING DNA CODING
; FOR MYOSIN HEAVY CHAIN SM1 ISOFORM PROTEIN INSERTED INTO
; TITLE OF INVENTION: VECTOR DNA, MICROORGANISM CARRYING THE RECOMBINANT DNA, AND
; TITLE OF INVENTION: AN AGENT FOR TREATMENT OF ARTERIOSCLEROSIS COMPRISING THE
; RECOMBINANT DNA
; FILE REFERENCE: 07898-013001
; CURRENT APPLICATION NUMBER: US/08/875,435B
; CURRENT FILING DATE: 1997-07-25
; PRIOR APPLICATION NUMBER: PCT/JP96/00134
; PRIOR FILING DATE: 1996-01-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 5
; LENGTH: 6644
; TYPE: DNA
; ORGANISM: Oryctolagus cuniculus
US-08-875-435B-5

```

```

Alignment Scores:
Pred. No.: 1,38e-12 Length: 6644
Score: 268.50 Matches: 178
Percent Similarity: 35.74% Conservative: 104
Best Local Similarity: 22.56% Mismatches: 213
Query Match: 9.64% Indels: 295
DB: 4 Gaps: 36

```

US-10-023-523-44 (1-546) x US-08-875-435B-5 (1-6644)

```

QY 4 GlnAspLysLysAsnGly---AlaAlaLysGlnSerAsnProLysSerSerProGlyGln 22
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3549 GAAGACGAGCTGGAGGACACAGCTGGACACCCAGCGCCACCCAGCAGAGCTCCGGGCCAA 3608
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 23 ProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAlaValGlu 42
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3609 GCGGAGCAGGAGTACGCTGCTGAAGAGGCGCTGACGAGAGACCCGCTGCCATCA 3668
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 43 AlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAlaArgThr 62
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3669 GGCCACAGTCCAGGA---GATGAGGAGAGAAACACACACAGGTGGTGGAGAGCTCAC--- 3722
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 63 AlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGluAspIle 82
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3723 -----GGACAGCTGGAACAGTTCAGAGGGCGCAAGGGGCAACCTCGACAAAGAC--- 3770
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 83 LeuSerThrTyrCysValAspAsnAsnGlnGly----- 93
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3771 -----CAACACAGACGCTGGAGAGGAGAGAACGACAGACCTGCCGCGGAGTGGC 3818
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 94 -----GlyProGlyGlu-----AspGlyAlaGlnGlyGluProAla----- 105
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3819 GGTCTCTGGCGCAGCCCAAGCAGGAGGTGGAGCACAAGAGAGAGAGCTGGAGGTGCAGCT 3878
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 105 ----- 105
Db 3879 GCAGGAGCTGCAGTCCAAAGTGCAGCAGTGGGAGCGCGCGCGGAGCTCAACAGACAA 3938
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 106 GluProGluAspAlaGluLysSer-----ArgThrTyrValAlaArgAsnGlyGluPro 123
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3939 GTCCACAGCTGCAGATGAAGTGGAGAGCTGACGGCGATGCTCGCGAGCCGAGGG 3998
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 124 Glu-----ProThrProValValAsn 130
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3999 GAAGGCCATCAAGCTGGCCAAAGGAGGTGGCTCCCTCGGGTCCAGCTCCAGGATACCCA 4058
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 131 GlyLysLysGluProSerLysGlyAspProAsnThrGluGluIleArgGlnSerAspGlu 150
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4059 GGA---GCTGCTCCAAAGAGAAACCCCGCAGAGAGCTCAACGTTGCCACCAAGCTCGCGCA 4115
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 151 ValGlyAspArgAsp----- 155
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4116 GCTGAGGAGCAGAGGAGAACAGCCTGCAGGAGCAGCTGGACGAGGAGATGGAGGCCAAGCA 4175
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 156 -----HisArgArgProGlnGlu-----LysLysLysAla 165
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4176 GAACCTGGAGCGCCACCATCTCCACCTGAAATCCAGCTCTCCGACTCAAGAGAGAGCT 4235
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 166 LysGlyLeu-Gly-----LysG 171
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4236 GCAGGACTTGCAGCACCGTGGAGTCTCTGGAGGAGGCAAGAGAGTTCACAGAGCA 4295
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 171 uIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAla 191
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4296 AATTGAGAGCCTCACCCACAGCATAC-----GAAGAGAGAGAGAGCTGC 4337
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 191 a-----LeuCysLysLysTyrAlaGluLeuLeuGl 201
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 4338 TTACGATTAACCTGAAAGACCAAGAACAGGCTTTCAGAGGAGCTCGAGCAGCTGGTCT 4397
Qy 201 uGluHisArgAsnSerGlnLysGlnMetLysLeuGlnLysLysGlnSerGln----- 219
Db 4398 AGACTTGGATTAACAGCGGCAACTGGTGTCCAACTTGGAAAGAACAGCAAGAAAGTTGCA 4457
Qy 220 -----LeuValGlnGlnLysAsp----- 225
Db 4458 TCAGTTGTTAGCCGAGAAAGAACATCTCTCCAAAGTATGGGTGAAGGAGCCGAGC 4517
Qy 226 -----HisLeuArgGlyGlnHisSerLysAlaVal-----LeuAlaArgSerLysLeuGln 242
Db 4518 CGAGGCTGAAGCAAGCAAGAAAGAAACCAAGGCTTGTCTCGCTCGGCGCTCGAGCA 4577
Qy 242 uSerLeu-----CysArgGlnLeuGlnArgHisAsnArgSerLeuLysGlnGlu----- 258
Db 4578 GGCTTGGAGGCCAAAGAGAGCTCGAGAGAACCAACAAATCTCAAGCCGAGATGGA 4637
Qy 259 -----GlyValGlnArgAlaArgGluGlnGluGlnLysArgLys 271
Db 4638 AGACCTCGTCAGCTCCAAAGCAGCTGGGCAAGACGTCATGCTGGAGAAGTCCAA 4697
Qy 271 s-----GluValThrSerHis 276
Db 4698 GCGGGCCCTGGAGACACAGATGGAGGATGAAGACGACGCTGGAAGAGCTAGAGGACGA 4757
Qy 276 sPheGlnValThrLeuAsn----- 282
Db 4758 GCTGAGGCCACCGAGGACCCCAAGCTGCGGTTGGAGTCAACATCGAGGCCCTCAAGCT 4817
Qy 283 -----AspIleGlnLeuGlnMetGlnHisAsnGlnuArgAsnSerLysLe 298
Db 4818 CCAGTTCGAGCGGATCTCCAGGCCCGGATGAGCAGACGAGGAGAGGAGGAGCT 4877
Qy 298 uArgGlnGluAsnMetGlu-----Le 305
Db 4878 GCAGAGCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4937
Qy 305 uAlaGluArgLeuLysLysLeuLeuGlu-----GlnTyrGlnLeu----- 318
Db 4938 GCGCGCGGACCAAGAGAGAGCTGGAGGGGACCTGAAAGACCTGGAGCTTCAGGCGGA 4997
Qy 319 -----ArgGlnuHisLysLysValPheLys-----His 329
Db 4998 CTCGCGCATCAAAGGGCGGAGAGCCTCAAGACGCTTCTGAAACTCGAGGCTCAGAT 5057
Qy 329 sLysAspLeuGlnGlnLeuValAspAlaLys----- 340
Db 5058 GAAGGACTTCAGAGAACTGGAGATGCCGCTGCTCCAGAGACGAGATCTTTGCCAC 5117
Qy 341 -----LeuGlnGlnAlaGlnGln 346
Db 5118 AGCCAGGAGAACAGAGAAAGCCAGAGTCTGGAGGCGAGACCTCATGCGACTACAAGA 5177
Qy 346 uMetLeuLysGlnuAlaGlnuArgHisGlnArg-----GluLysAspPheLeuLe 363
Db 5178 GATCTCGCGCGGACAGAGGCTCCAAACAGGCGAGACTTCGAGAGAGGAGAGTGGC 5237
Qy 363 uLysGluAlaValGlnSerGlnArgMetCysGluLeuMetLysGlnGlnGlnuThrHisLe 383
Db 5238 CGAGGAGCTGCCAGCAGCTTGTCCGGAAGGAGCGCTCGAGATGAGAAGCGCGCT 5297
Qy 383 uLysGlnGlnLeuAlaLeuTyrThrGluLysPheGlnuGlnPheGlnAsnThrLeuSerLy 403
Db 5298 GGAGGCCCGGATCGCACAGCTGGAGGAGGAGCTGGAGGAGAACAGGGCAACATGAGGC 5357
Qy 403 s-----SerSerGluValPh 408
Db 5358 AATGAGCGACCGCTCCGCAAGGCTAGCAGCAGCGCGAGCAGCTCAGCAACGAGCTGGC 5417
Qy 408 eThr-----ThrPheLysGlnGlnMetGlnLysMetTh 419
Qy 419 ----- 419

Db 5418 CACAGAGCGCAGCAGCAGCCAGAAAGATAGAGCGCAGCGCAGCAGCTCGAGCGCGAGAA 5477
Qy 419 rLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 438
Db 5478 CAAGAGCTCAAGAGCAAGCTGCGAGGAGATGGAGGGGCGAGTCAAGTCCAAAGTTCAGGTC 5537
Qy 438 rSerAsnLysAlaLeu-----LeuGluMetAl 447
Db 5538 CACTATCGCGCGCTGGAGGCCAAGATTGCGCAGCTGGAGGAGCAGGTTGAGCAGGAGGC 5597
Qy 447 aGluLysLysThrVal-----ArgAspLysGlnLeuGluGlyLe 460
Db 5598 CAGAGAGAGCAGCGCGCGCGCCCAAGCGCTGAAGCAGAGGAGGAGCTGAAGGAGAT 5657
Qy 460 uGlnValLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 477
Db 5658 GCTGCTGAGCTGGAAGAGCAGCGCAAGATGGCTGA-GCAGTACAAAGGAGCAGGAGGAG 5716
Qy 477 uArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlnGlnGlnGlnGlnGln 497
Db 5717 AAGGAAACGCCAAGCTCAAGCAGCTCAAGAGGAGCTGGAGGAGGCGGAGGAGGAGTCC-- 5774
Qy 497 rAspSerGlyProGluArgArgProGlu-----GlyProGln 509
Db 5775 -----GCAGCGCATCAACGCCCAAGCAGGAGCTGCAGCGGAGCTGCAGCGGCGCGG 5830
Qy 509 yAlaGlnAlaProSerSerProArg 517
Db 5831 AGAGCAACGAGCGCATGGCGCGCGA 5855
RESULT 26
US-09-976-594-640
; Sequence 640, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 640
; LENGTH: 5185
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 346209.3
US-09-976-594-640
Alignment Scores:
Pred. No.: 3,766-12 Length: 5185
Score: 261.50 Matches: 145
Percent Similarity: 37.85% Conservative: 99
Best Local Similarity: 22.52% Mismatches: 235
Query Match: 9.39% Indels: 167
DB: 4 Gaps: 26
US-10-023-523-44 (1-546) x US-09-976-594-640 (1-5185)
Qy 2 LysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSer----- 18
Db 542 CGAGAGAGGCGCGGAGGAGCTGGAGCGCGCTGGTGTGGACACGAGCTTCGCTGCAACGG 601
Qy 19 -----SerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAla 37
Db 602 GGTCTGGAGGCGCATCGCATCTGTCGCGCAGGGCTTCCCAACCGCATCTCTCTCCAGGA 661
Qy 38 AlaProAlaVal----- 41

Alignment Scores:

Pred. No.: 1,17e-11 Length: 6617
Score: 257.50 Matches: 147
Percent Similarity: 41.15% Conservative: 111
Best local Similarity: 23.44% Mismatches: 181
Query Match: 9.25% Indels: 191
DB: 4 Gaps: 31

US-10-023-523-44 (1-546) x US-09-976-594-268 (1-6617)

QY 6 LysLysAsnGlyAlaAlaLysGln-----SerAsnProLysSerSerProGlyGlnPro 23
DB 765 AAGAGTCAAGTGGAGCTGAGAGCAAGAGGCTCCGAGTCTGCAGCAGCAGCAGACCA 824
QY 24 GluAlaGly-----ProGluGlyAla----- 30
DB 825 GTAGCTGAGGCCCCAGTCCCAAGAGCTTGGATGTGCGGACAAGCAGGCTGGACTCCC 884
QY 31 GlnGluA-gProSerGlnAlaAlaProAlaVal-----GluAlaGluGlyProGly 47
DB 885 CAGGAGCACCAGGCTTGAGTGGAGAAGCTGTGTGTACAGAGAGCGCGCAGGA---GGA 941
QY 48 SerSerGlnAlaPro-----ArgLysProGluGly----- 57
DB 942 GCAGGAGAGGCTGCATGCCATCTCTTCGCGCTCGGAGAACAGGAGCTAAACATCACC 1001
QY 57 ----- 57
DB 1002 ATCATCTAAGCGGGTCAAGGAATTGAAAAAACAACATTTAAGGGTTAATA 1061
QY 58 -----AlaGlnAlaArgThrAlaGlnSer 65
DB 1062 TCCTACACAAATTCATTTACTTCATTTGAATGTAGAGCCCGCAGCAGCCATCTAAGGTA 1121
QY 66 GlyAlaLeuArgAspValSerGlnLeuSerArgGlnLeuGluAspIleLeuSerThr 85
DB 1122 GAGCAGCTACAGATGACAGCAACCACTAAGAGAGAGCTAGAGAGTGTGGG-AAGACA 1180
QY 86 TyrCysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAlaGlnGlyGluProAla 105
DB 1181 -----GCTCAGGCTGAGGTGGAACAAATCAGATGTTGAG 1216
QY 106 GluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsnGlyGluProGluPro 125
DB 1217 TCTCTCTGAACA-CGAGACAGGAGGAGAGGCTACGTACGTGAACAGGAGG----- 1260
QY 126 ThrProValValAsnGlyLysGluProSerLysGlyAspProAsnThrGluGluLe 145
DB 1261 -----AGAGGCTACGTGAACAG-CAGGAGAGG 1286
QY 146 ArgGlnSerAspGluValGlyAspArgAspHisArgArgProGlnGlu-----LysLysLys 164
DB 1287 CAACGTGAACAGAG-----GATAGGCTACATGAACAGAGGAGGAGGCTACGTGAACAG 1340
QY 165 AlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThr 184
DB 1341 GAGGAGAGGCTGTGTGAACAGGAGGAGGCTGTGTGAACAGGAGGAGGCTACGTGAA 1400
QY 185 ProGluGluLysLeuAla-----AlaLeuCysLysLysTyrAlaGluLeuLeu 200
DB 1401 CATGAGGAGGCTGTGTGAACAGGAGGAGGAGGCTATGTGAACAGGAGGAGGCTACGT 1460
QY 201 GluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysSerGlnLeu 220
DB 1461 GAACAGGAGGAGGCTACATGAACAGGAGGAGGCTACGTGAACAGGAGGAGGCTG 1520
QY 221 ValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLys 240
DB 1521 TGTGAACAGGAGGAGGCTACGT---GAACATGAG----- 1553
QY 241 LeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGluGlyVal 260
DB ----- 260

RESULT 28

US-08-466-390-3
; Sequence 3, Application US/08466390
; Patent No. 5686562
; GENERAL INFORMATION:
; APPLICANT: TONKATLY, GARY
; APPLICANT: LIDGARD, GRAHAM P

TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE
 TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: TESTA, HURWITZ & THIBEAULT

STREET: 125 HIGH STREET

CITY: BOSTON

STATE: MA

COUNTRY: USA

ZIP: 02110

COMPUTER READABLE FORM: disk

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/466,390

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: PITCHER ESO, EDMUND R

REGISTRATION NUMBER: 27,829

REFERENCE/DOCKET NUMBER: MTP-013

TELEPHONE: (617) 248-7000

TELEFAX: (617) 248-7100

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 6306 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 1-6306

PUBLICATION INFORMATION:

AUTHORS: COMPTON, DUANE A

AUTHORS: SZILAK, ILYA

AUTHORS: CLEVELAND, DON W

TITLE: PRIMARY STRUCTURE OF NUMA, AN INTRANUCLEAR

TITLE: PROTEIN THAT DEFINES A NOVEL PATHWAY FOR

TITLE: SEGREGATION OF PROTEINS AT MITOSIS

JOURNAL: J. Cell Biol.

VOLUME: 116

PAGES: 1395-1408

DATE: 1992

US-08-466-390-3

Alignment Scores:
 Pred. No.: 1,46e-11 Length: 6306
 Score: 256.00 Matches: 155
 Percent Similarity: 36.84% Conservative: 108
 Best Local Similarity: 21.74% Mismatches: 254
 Query Match: 9.19% Indels: 198
 DB: 1 Gaps: 27

US-10-023-523-44 (1-546) x US-08-466-390-3 (1-6306)

QY 6 LysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerProGlyGlnProGluAla 25
 Db 1587 AGAAGAGCGAGCGCCAGCTAGCAGACCCCTCCAAAGCAAGCAAGAGGCGCTCCAGGCGCT 1646
 QY 26 GlyProGluGly-----AlaGlnGluArgProSerGlnAlaAlaProAlaValGluAla 43
 Db 1647 CCGCCACCGAGGTGGAGCAGCTAGCAGTAGCTGAGCAGAGGAGGAGCAGCTTGAAGGA 1706
 QY 44 GluGlyProGlySerSer-GlnAlaProArgLysProGluGlyAlaGlnAlaArgThrAl 63
 Db 1707 ---GGTAGCGGAGAGCAGGAGGCACTAGGCGAGGACCATGCCAGCACTGGCCACTGC 1763
 QY 63 a-----GlnSerGlyAlaLeuArgAsp-----70

Db 1764 TGCAGAGGAGCGAGAGCGCTCCCTTAAGGGAGCGGATGCGGCTCTCAAGCAGCTGCGGC 1823
 QY 71 -valSerGluGluLeuSerArgGlnLeuGluAspPileLeuSerThrTyrcysValaspas 90
 Db 1824 ACTGAGAGGAGGAGGAGGCTCCCAAGCTGGAGATTCTGCAGCAGCAACTTCCAGTGGCTAA 1883
 QY 90 nasGlnGlyGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluProGluAspAl 110
 Db 1884 TGAA-----GCCCGGAGCAGTGCACAGCTCAGTCACACAG-----GC 1922
 QY 110 aGlySerArgThrTyValalaArgAsnGlyGluProGluProThrProValValas 130
 Db 1923 CCAGCGGAGAGGAGCAGAGCTGAGCCGAGAGGTGGAGGAAGCTCCAGGCGCTGTGTGAGAC 1982
 QY 130 nGlyGluGlyProSerLysGlyAspProAsnThrGluGlu-----IleAr 146
 Db 1983 AGCCCGCAGMACATGAGCCCGCAGGCTTGCAGAGCTAGAGTTGCAGCTGG 2042
 QY 146 gGlnSerAspGluValGlyAspArgAspHisArgArgProGlnGluLysLysAla 166
 Db 2043 GTCTGAGCAGCAAAAGCAACTCAGAAAGAAAGGTGGCCAGGAGAGGACAG----- 2097
 QY 166 sGlyLeuGlyLysGluLeuThrLeuMetGlnThrLeuAsnThrLeu-----Serth 184
 Db 2098 ----CTCCAGGAGCAGCTCAGGCGCTCAAAGAGTCTTGAAGTCAACAAGGCGAGCCT 2153
 QY 184 rProGluGluLys-----LeuAlaAl 191
 Db 2154 TGAAGAGAGAGCGCAGGCGCTGCAGATGCCCTGGAGAGCAGCAGCGTTGATCTCTGA 2213
 QY 191 aLeuCysLysTyAlaGluLeuGluHisArgAsnSerGlnLysGlnMet--- 210
 Db 2214 GCTGAAGGAGAGAGCGGAGCGCTGTGGAGCAGCATAGCGGAGCAAGAGAGAGCTGA 2273
 QY 211 -----LysLeuLeuGlnLysGlnSerGlnLeuValGln----- 222
 Db 2274 AGAAGAGAGGCGTGGCGGCGCAGGCGTGGAGGCTCGATTACTGCAGCTTGGGGAGGCCCA 2333
 QY 223 -----GluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerly 240
 Db 2334 TCAGGCTGAGACTGAAGTCTCGCGCGGAGGCTGGCAGAGGCGCATGGCTGCCAGCACAC 2393
 QY 240 sLeuGluSerLeuCysArgGluLeu----- 248
 Db 2394 AGCTGAGAGTGTGAGCAGCTCTCAAAGAGTAGTGTGCGCTGCGGTGACGGGTATGA 2453
 QY 249 -----GlnArgHisAsnArgSerLe 255
 Db 2454 GGATAGCCAGAGAGGAGGAGCAGTATGGCGGCGCATGTTCCAGGAACAGCTGATGACTTT 2513
 QY 255 uLysGluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGlu----- 272
 Db 2514 GAAGGAGGAA---TGTGAGAGGCGCCCGCAGGAGCTGCAGGAGGAGGAGAGAGGTGGC 2570
 QY 273 ----ValThrSerHisPheGlnValThrLeuAsnAspPileGln----- 285
 Db 2571 AGGCATAGATCCACAGCAGGCTCCAGATAGCGCGCAGCAGCAACAACTAGCTGAGCT 2630
 QY 286 -----LeuGlnMetGluGlnHisAsnGlu---ArgAsnSerly 297
 Db 2631 CCATGCCAACCTGGCCAGAGCAGCTCCAGCAGGTCCAAGAGAGAGAGTCAAGGCGCCAGAA 2690
 QY 297 sLeuArgGlnGlnAsnMetGluLeuAlaGlu----- 307
 Db 2691 GCTTGCAGATGACTCTCCACTCTGCAGGAAGAGTGGTCCGCCAGCAAGAGGTGGC 2750
 QY 308 -ArgLeuLysLysLeuLe----- 313
 Db 2751 CCGCTTGGAGACCTTGGTCCCGAAGGAGCAGGTGAGCAGCAGGAAACAGCCTCCCGGAGTT 2810
 QY 314 -----GluGlnTyGluLeuArgGluGluHisIleAs 324
 Db 2811 AGTCAAGGAGCGCTGCGAGGCGCAGGAGCAGACAGCGCCGAGTGGCTGGAAGAGCAACAGGG 2870

CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: PITCHER ESQ. EDMUND R
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: WTP-013
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 6306 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..6306
PUBLICATION INFORMATION:
AUTHORS: COMPTON, DUANE A
AUTHORS: SZILAK, ILYA
AUTHORS: CLEVELAND, DON W
TITLE: PRIMARY STRUCTURE OF NIMA, AN INTRANUCLEAR
TITLE: PROTEIN THAT DEFINES A NOVEL PATHWAY FOR
TITLE: SEGREGATION OF PROTEINS AT MITOSIS
JOURNAL: J. Cell Biol.
VOLUME: 116
PAGES: 1395-1408
DATE: 1992
US-08-457-781-3

Alignment Scores:

Pred. No.:	1.46e-11	Length:	6306
Score:	256.00	Matches:	155
Percent Similarity:	36.8%	Conservative:	108
Best Local Similarity:	21.7%	Mismatches:	254
Query Match:	9.19%	Indels:	198
DB:	1	Gaps:	27

US-10-023-523-44 (1-546) X US-08-467-781-3 (1-6306)

Qy	6	LysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerProGlyGlnProGluAla	25
Db	1587	AGAGAAGCAGGCCCTAGCACAGACCTCCAAACAGAGAAGCAGGCTCCACAGGCGCT	1646
Qy	26	GlyProGluGly-----AlaGlnGluArgProSerGlnAlaAlaProAlaValGluAla	43
Db	1647	CCGCCACAGTGGAGAGCAGCTAGACAGTAGCTGAAGCAGAGGAGGAGCAGGTGAAGA	1706
Qy	44	GluGlyProGlySerSer-GlnAlaProAlaArgLysProGluGlyAlaGlnAlaArgThrAl	63
Db	1707	--GCTAGCGAGAAGCAGGAGGCAACTAGGACGAGCATGCCAGCAACTGGGCACTGC	1763
Qy	63	a-----GlnSerGlyAlaLeuArgAsp-----	70
Db	1764	TGCAGAGGAGCAGAGGCGCTCTTTAAGGAGCGGGATCGCGCTCTCAAGCAGCTGGAGGC	1823
Qy	71	-ValSerGluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAs	90
Db	1824	ACTGGAGAAGGAGAAGGCTCCCAAGCTGGAGATTCTGCACGACCAACTTCAGGTGGCTAA	1883
Qy	90	nAsnGlnGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluProGluAspAl	110
Db	1884	TGAA-----GCCCGGAGACAGTGGCCAGACCTCAGTGACACAG-----GC	1922
Qy	110	aGluLysSerArgThrTyrValAlaArgAsnGlyGluProGluProThrProValValAs	130
Db	1923	CCAGCGGAGAAGGCAGACAGCTGAGCCGGAAGTGGAGAACTCCAGGCTGTGTTGAGAC	1982
Qy	130	nGlyGluLysGluProSerLysGlyAspProAsnThrGluGlu-----IleAr	146
Db	1983	AGCCCGCAGGACAGCATGAGCCAGGCCAGCCAGGTTGCAGAGTACAGTTGCAGGTGCG	2042

Qy	146	gGlnSerAspGluValGlyAspArgAspHisArgProGlnGluLysLysAlaLys	166
Db	2043	GTCTGACGACAAAAAGCAACTCAGAAAGAAAGGGTGGCCACAGAGAAGCACAG	2097
Qy	166	scGlyLeuGlyGluLeuThrLeuLeuMetGlnThrLeuAsnThrLeu	184
Db	2098	CTCCAGAGCAGCTCCAGGCCCTCAAGAGTCTCTTAAGGTCAACAGGCGACGCT	2153
Qy	184	rProGluGluLys	191
Db	2154	TGAAGAGGAGAACGGCAGGGCTCGAGATGCCCTGGGAAGACGACGAGCGTTGTATCTCTGA	2213
Qy	191	aLeuCysLysLysTyraLaGluLeuLeuGluHisArgAsnSerGlnLysGlnMet	210
Db	2214	GCTGAAGGCAGACACCGCAAGCCCTGTGGAGGAGCATAAAGCGGGAACGAAGAGAGTGGGA	2273
Qy	211	LysLeuLeuGlnLysLysGlnSerGlnLeuValGln	222
Db	2274	AGAGAGAGGGCTGGGGCGCAAGGGCTGGAGGCTCGATTACTCGCAGCTTGGGAGGCCCCA	2333
Qy	223	GluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLys	240
Db	2334	TCAGGCTGACACTCAAGTCTCGCGGGGGAGCTGGCAGAGGCCATGGCTGCCAGCACAC	2393
Qy	240	sLeuGluSerLeuCysArgGluLeu	248
Db	2394	AGCTGAGAGTCACTGTGAGCAGCTCTGTCAAGAAGTAGCTGCTGGCGTGACGGGTATGA	2453
Qy	249	GlnArgHisAsnArgSerLeu	255
Db	2454	GGATAGCCAGCAAGAGGAGGCACAGTATGGCGGCATGTTCCAGGAACAGCTGATGACTTT	2513
Qy	255	uLysGluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGlu	272
Db	2514	GAAGGAGGAA--TGTGAGAGGCCCGCCAGGAGCTGCAGAGGCAAGAGAGGTGGC	2570
Qy	273	ValThrSerHisPheGlnValThrLeuAsnAspIleGln	285
Db	2571	AGGCATAGAATCCACAGCAGAGCTCCAGATAAGCGCGCAGCAGACAACTAGCTGAGCT	2630
Qy	286	LeuGlnMetGluGlnHisAsnGlu	297
Db	2631	CCATGCCAACCTGGCCAGAGCACTCCAGCAGGTCCCAAGAGAAGAAAGTCAAGGCCACAGAA	2690
Qy	297	sLeuArgGlnGluAsnMetGluLeuAlaGlu	307
Db	2691	GCITGCAGATGACCTCTCCACTCTGCAGGAAGAAGATGGCTGCCACCAGCAAGAGGTGC	2750
Qy	308	ArgLeuLysLysLeuLeu	313
Db	2751	CCGCTTGGAGACCTTGGTCCGCAAGGACGTGAGCAGAGGAACAACGCTCCCGGAGTT	2810
Qy	314	GluGlnTyrGluLeuArgGluGluHisIleAs	324
Db	2811	AGTCAAGGACCTGCCAGGGCAGGACACACAGCCCGAGTGGCTGTAAGACCAACAGGG	2870
Qy	324	pLysValPheLysHisLysAspLeuGlnGlnLeuValAspAlaLysLeuGlnAla	344
Db	2871	ACGCCAGTTCTGCAGCACACAGGCAGCGCTGCAGGCTATGGAGCGGGAGCAGACAGAT	2930
Qy	344	aGlnGluMetLeu	352
Db	2931	GGGCATGAGCTGGAACGGCTCGGGGCCCGGCTGTGTGAGAGCCAGGGCGCAGCAGGGA	2990
Qy	352	uGluArgHisGlnArgGluLysAspPhe	370
Db	2991	GGAGCGTGGCAGCAGCAAGGAGGTGGCGGGCTGACCCAGGAGCGGGCGCGTGCCTCA	3050
Qy	370	nArgMetCysGluLeuMetLysGlnGlnGluThrHisLeuLysGlnLeuAlaLeuTy	390
Db	3051	GGCTGACCTTGGCCCTGGAGAGCGCGCCACAGACAGAGCTTGAATGCGGCTG	3102
Qy	390	rThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluValPheThr	410

FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..6306
 PUBLICATION: INFORMATION:
 AUTHORS: COMPTON, DUANE A
 AUTHORS: SZILAK, ILYA
 AUTHORS: CLEVELAND, DON W
 TITLE: PRIMARY STRUCTURE OF NUMA, AN INTRANUCLEAR
 TITLE: PROTEIN THAT DEFINES A NOVEL PATHWAY FOR
 TITLE: SEGREGATION OF PROTEINS AT MITOSIS
 JOURNAL: J. Cell Biol.
 VOLUME: 116
 PAGES: 1395-1408
 DATE: 1992
 US-08-483-924-3

Alignment Scores:
 Pred. No.: 1.46e-11 Length: 6306
 Score: 256.00 Matches: 155
 Percent Similarity: 36.83% Conservative: 108
 Best Local Similarity: 21.71% Mismatches: 254
 Query Match: 9.19% Indels: 198
 DB: 2 Gaps: 27

US-10-023-523-44 (1-546) x US-08-483-924-3 (1-6306)

QY 6 LysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerProGlyGlnProGluAla 25
 Db 1587 AGAAGACGCGCCAGCTAGCAGACCCCTCCAAACAGCAAGAACAGGCTCCAGGGCT 1646
 QY 26 GlyProGluGly-----AlaGlnGluArgProSerGlnAlaAlaProAlaValGluAla 43
 Db 1647 CCGCCACCGAGTGGAGCAGCTAAGCTAGCTAGCTTGAAGCAGAGGACGACGCTTGAAGA 1706
 QY 44 GluGlyProGlySerSer-GlnAlaProArgLysProGluGlyAlaGlnAlaArgThrAl 63
 Db 1707 ---GGTAGCGGAGAGCAGGAGGCACTAGGAGGACCATGCCAGCACTGGCCACTGC 1763
 QY 63 a-----GlnSerGlyAlaLeuArgAsp----- 70
 Db 1764 TGCAGAGCGGAGAGGCGCTCTTAAGGAGCGGATGCGGCTCTCAAGCAGCTGGAGGC 1823
 QY 71 -ValSerGluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAs 90
 Db 1824 ACTGGAGAGGAGAGGCTCCAGCTGGAGATCTGCAGCAGCACTTCAGGTGCTTAA 1883
 QY 90 nAsnGlnGlyGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluProGluAspAl 110
 Db 1884 TGAA-----GCCCGGAGCAGTGGCCAGACCTCAGTGACACAG-----GC 1922
 QY 110 aGluLysSerArgThrTyrValAlaArgAsnGlyGluProGluProThrProValValAs 130
 Db 1923 CCAGCGGAGAGAGCAGAGCTAGCGCCGAGAGGTGGAGAACTCCAGCCCTGTGTGAGAC 1982
 QY 130 nGlyGluLysGluProSerLysGlyAspProAsnThrGluGlu-----IleAr 146
 Db 1983 AGCCCGCCAGGACAGCATGAGCCCGAGGCGCCAGGTTGCAGACTAGAGTTGAGCTGCG 2042
 QY 146 gGlnSerAspGluValGlyAspArgAspHisArgArgProGlnGluLysLysLysAlaL 166
 Db 2043 GTCTGAGCAGCAAAAGCAACTGAGAAAGAAAGGTGGCCCGGAGAGAGGACACAG----- 2097
 QY 166 sGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeu-----SerTh 184
 Db 2098 ----CTCAGAGAGCAGCTCCAGGCCCTCAAGAGTCTCTGAAGGTCCACCAAGGCGAGCT 2153
 QY 184 rProGluGluLys-----LeuAlaAl 191
 Db 2154 TGAAGAGAGAGAGCGGCGGCTGCAGATGCCCTGGAGAGCAGCAGCGTGTGTATCTCTGA 2213
 QY 191 aLeuCysLysLysTyrAlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMet-- 210
 Db 2214 GCTGAAGCGAGAGACCCGAGCCTGTGTGGAGCAGCATAGCGGAGACGAGAGGAGCTGGA 2273

QY 211 -----LysLeuLeuGlnLysLysGlnSerGlnLeuValGln----- 222
 Db 2274 AGAAGAGAGCGCTGGCGCAAGGGCTGGAGGCTCGATTACTGACGCTTGGGAGAGCCCA 2333
 QY 223 -----GluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLy 240
 Db 2334 TCAGGCTGAGACTCAAGCTCTCGCGCGGAGCTGGCAGAGCCATGGTGCACAGCACAC 2393
 QY 240 sLeuGluSerLeuCysArgGluLeu----- 248
 Db 2394 AGCTGAGAGTGTGTGAGCAGCTCGTCAAAGAAGTAGTCTGCTGCTGACGGGTATGA 2453
 QY 249 -----GlnArgHisAsnArgSerLe 255
 Db 2454 GGATAGCCAGCAAGAGGAGCAGTATGGCGCATGTTCCAGGAAACAGCTGATGACTTT 2513
 QY 255 uLysGluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGlu----- 272
 Db 2514 GAAGGAGGAA---TGTGAGAGAGCCCGCCAGAGAGCTGCAGAGGCAAGAGAGAGTGGC 2570
 QY 273 ----ValThrSerHisPheGlnValThrLeuAsnAspIleGln----- 285
 Db 2571 AGGCATAGAAATCCACAGCGAGCTCCAGATAAGCCGACAGCAACAACTAGCTAGCT 2630
 QY 286 -----LeuGlnMetGluGlnHisAsnGlu---ArgAsnSerLy 297
 Db 2631 CCATGCCAACCTGCCAGAGCAGCTCCAGAGAGTCCAGAGAGAGAGTCCAGGCCAGAA 2690
 QY 297 sLeuArgGlnGluAsnMetGluLeuAlaGlu----- 307
 Db 2691 GCTTCAGATGACCTCTCCACTCTGCAGGAAAGATGGCTGCCACAGCAAGAGGTGGC 2750
 QY 308 -ArgLeuLysLysLeuLe----- 313
 Db 2751 CCGCTTGAGACCTTGGTCCGCAAGGCGAGTGCAGCAGCAAGAACAGCTCCCGGAGTT 2810
 QY 314 -----GluGlnTyrGluLeuArgGluGluHisLeAs 324
 Db 2811 AGTCAAGGAGCTCGAGGGCAGGAGACAGACCCGAGTGGCTGGAGAGCAACAGGG 2870
 QY 324 pLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnAl 344
 Db 2871 ACGCCAGTTCAGCAGCACACAGCAGCGCTGCAGGCTATGGAGCGGAGGAGCAGAGAT 2930
 QY 344 aGlnGluMetLeu-----LysGluAlaGl 352
 Db 2931 GGGCAATGAGCTGGAAACGGCTGCGGCGCGCTGATGGAGAGCCAGGCGGAGCAGCAGGA 2990
 QY 352 uGluArgHisGlnArgGluLysAspPhe-----LeuLeuLysGluAlaValGluSerGl 370
 Db 2991 CGAGCGTGGCAGCAGAGAAAGAGGTGGCGGCTGACCCAGGAGCGGCGCGCTGCCCA 3050
 QY 370 nArgMetCysGluLeuMetLysGlnGlnGluThrHisLysLysGlnLeuAlaLeuLy 390
 Db 3051 GGCTGACCTTCCTCGAGAGAGGGCGGCGCAGCAGAGCTTGAGATGCGGCTG----- 3102
 QY 390 rThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluValPheThrTh 410
 Db 3103 -----CAGAACGCCCTCAACGAGCAGCGCTGTGGAGTTCTGGTAC 3140
 QY 410 rPheLysGln----- 413
 Db 3141 CTTGCAAGAGGAGCAGTGGCTCATGCCCTGACGGAAAGAGAGGCAAGCAGGAGTTGGC 3200
 QY 414 -----GluMetGluLysMetThrLysLysIl 422
 Db 3201 CAAGCTTCGTGTCTGGAGGAGCCCGCATAAAGAGCTGGAGGAACTTCGGCAACCGCT 3260
 QY 422 eLysLysLeuGluLysGluThrThrMetTyrArgSerArgTrpGluSerSerAsnLysAl 442
 Db 3261 GAAGCAACTGAAGGAACACAGCTGGCTTAAGAAAGAAAGAGGACCGCATCTGCTCAGGAGC 3320

QY 442 aLeuLeuGluMetAlaGluLysThrValArgAspLysGluLeuGluGlyLeuGlnVa 462
 Db 3321 CCAATCTGAGGCTGCTGGCAGACAGACAGCCACAGCGCCCAAGCTGGAGCACTCGGGC 3380
 QY 462 llystleGlnArgLeuGluLysLeuGluGlnThrValArgAspLysGluLeuGlnVa 482
 Db 3381 AGAGGTGAGCAAGCTGGAACAGCAATGCGAAGACAGCAGGAGCGGTGACAGCCTGGA 3440
 QY 482 nlystleGlnArgLeuGluLysLeuGluGlnThrValArgAspLysGluLeuGlnVa 498
 Db 3441 AGCAGCTCGAGGCTGAGCGGCTGCGGGCTGGA-GCGGACAGCTGCTCGGAGACTC 3499
 QY 498 pSerGlyProGluArgArgProGluGlyProGluGlnAlaProSer-----Se 515
 Db 3500 TGCAGGGCCAGTT---AGAGGAGAGGCGCCAGGAGCTAGGCGACAGTCAGAGTGCCTTAG 3556
 QY 515 rProArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAlaSerGly-- 534
 Db 3557 CTTGCGGCCCAACGGGAGTTGGCTGCTTCCGACCAAGTACAGACCCACAGCAGGCTG 3616
 QY 535 -----GlnThrGlyProGlnGluProThrSerAlaArg 545
 Db 3617 AGATGAGTGGAGGCCCAAGTGGCGCGGGCGGCGCAAGA 3656

RESULT 33

PCT-US93-06160-3

; Sequence 3, Application PC/TUS9306160

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE

; TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESSES:

; ADDRESSEE: TESTA HURWITZ & THIBEAULT

; STREET: 53 STATE STREET

; CITY: BOSTON

; STATE: MA

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US93/06160

; FILING DATE: 19930621

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: PITCHER ESQ, EDMUND R

; REGISTRATION NUMBER: 27,829

; REFERENCE/DOCKET NUMBER: MTP-013

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617/248-7000

; TELEFAX: 617/248-7100

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 6306 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 1..6306

; PUBLICATION INFORMATION:

; AUTHORS: COMPTON, DUANE A

; AUTHORS: SZILAK, ILLYA

; AUTHORS: CLEVELAND, DON W

; TITLE: PRIMARY STRUCTURE OF OF NUMA, AN INTRANUCLEAR

; TITLE: PROTEIN THAT DEFINES A NOVEL PATHWAY FOR

; TITLE: SEGREGATION OF PROTEINS AT MITOSIS

; JOURNAL: JOURNAL OF CELL BIOLOGY

; VOLUME: 116
 ; PAGES: 1395-1408
 ; DATE: MAR-1992
 PCT-US93-06160-3

Alignment Scores:

Pred. No.: 1 46e-11 Length: 6306
 Score: 256.00 Matches: 155
 Percent Similarity: 36.83% Conservative: 108
 Best Local Similarity: 21.71% Mismatches: 254
 Query Match: 9.19% Indels: 198
 DB: 5 Gaps: 27

US-10-023-523-44 (1-546) x PCT-US93-06160-3 (1-6306)

QY 6 LysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerProGlyGlnProGluAla 25
 Db 1587 AGAAGACAGCGCCAGCTAGCAGACAGCCCTCCACACAGCAAGAACAGCGCTCCAGGGCCT 1646
 QY 26 GlyProGluGly-----AlaGlnGluArgProSerGlnAlaAlaProAlaValGluAla 43
 Db 1647 CCGCCACCAGTGGAGCAGCTAAGCAGTAGCTCCGAGCAGGAGGAGGAGGAGGAGGAGGAG 1706
 QY 44 GluGlyProGlySerSer-GlnAlaProArgLysProGluGlyAlaGlnAlaArgThrAl 63
 Db 1707 ---GCTAGCGAGAGAGCAGGAGGCAACTAGCAGGACCATGCCAGCACTGGCCACTGC 1763
 QY 63 a-----GlnSerGlyAlaLeuArgAsp----- 70
 Db 1764 TGCAGAGGAGCAGAGGCGCTCTTAAGGGAGCGGGATGCGGCTCTCAAGCAGCTGGAGGC 1823
 QY 71 -ValSerGluGluLeuSerArgGlnLeuGluAspLleLeuSerThrTyrCysValAspAs 90
 Db 1824 ACTGGAGAGGAGAGGCTGCCAGCTGGAGATTCTGCAGCAGCAACTTCAGGTGGCTAA 1883
 QY 90 nAsnGlnGlyGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluProGluAspAl 110
 Db 1884 TGAA-----GCCCGGACAGCTGCCAGACCTCAGTCACACAG-----GC 1922
 QY 110 aGluLysSerArgThrTyrValAlaAlaArgAsnGlyGluProGluProThrProValValas 130
 Db 1923 CCAGCGGAGAGGAGCAGAGCTGAGCCGGAAGGTGGAGGAACCTCCAGGCTGTGTTGAGAC 1982
 QY 130 nGlyLysGluProSerLysGlyAspProAsnThrGluGlu-----lleAr 146
 Db 1983 AGCCCGCCAGGAACAGCATGAGGCCAGCCAGGTTGCAGAGCTAGAGTTGCAGCTGCG 2042
 QY 146 gGlnSerAspGluValGlyAspArgAspHisArgArgProGlnGluLysLysLysAlaLy 166
 Db 2043 GTCTGAGCAGCAAAAGAGCACTGAGAAAGAGGGTGGCCCGCCAGGAGAGGAGGAGGAGGAG 2097
 QY 166 sGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeu-----SerTh 184
 Db 2098 ---CTCCAGGAGCAGCTCCAGGCGCTCAAGAGTCTCTGAAGGTCAACAGGGCAGCCT 2153
 QY 184 rProGluGluLys-----LeuAlaAl 191
 Db 2154 TGAAGAGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2213
 QY 191 aLeuCysLysLysTyrAlaGluLeuLeuGluHisArgAsnSerGlnLysGlnMet-- 210
 Db 2214 GCTGAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2273
 QY 211 -----LysLeuLeuGlnLysLysGlnSerGlnLeuValGln----- 222
 Db 2274 AGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2333
 QY 223 -----GluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLy 240
 Db 2334 TCAGGCTGAGACTGAAGTCTCTGCGGGGAGGAGTGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAG 2393
 QY 240 sLeuGluSerLeuCysArgGluLeu----- 248

1360 ---ACAAGAGCATCAGCTCCACAGTCTCCACCTCAAGTGTCTCGTCACAAATCACATTAT 1416
137 LysGlyAspProAsnThrGluGluLeuLeuArgGlnSerAspGluValGlyAspArgAspHis 156
1417 CGTAATAGAGAAACACTTTGCAACTATACGAACAGCATCACTGGTT---ACA 1464
157 ArgArgProGlnGluLysLysLysAlaLysGlyLeuGlyLysGluLeuLeuLeuLeuLeuLeu 173
1465 AGCAGATCGAAGACATCAGCAGAGGACTCTGACTTAGAGACAGATCTCTGTTTATAG 1524
174 ---LeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeu 192
1525 CGGATCAGGCGACAGCATCAGAGCAGCTGATGACTCTCGAAATAAATACTGAAGGCAGAA 1584
193 CysLysLysTyraAla-----GluLeuLeuGluGluHisArgAsn-----Ser 206
1585 ATGACGAACATCGGCTCAGATTAGCAAGATCTTTGAACACTCAGCGCAACAAATTCGCT 1644
207 GlnLysGlnMetLysLeuGlnLysLysGlnSerGlnLeuValGlnGlu----- 223
1645 GCAGAAATCGAGAACTTATTAGAAACACCAAGCTTCTGTGAAATAAAGAGGCTAAAGTG 1704
224 -----LysAspHisLeuArgGlyGluHisSerLysLysAlaVal 235
1705 ATGGCCACAGGAGAGAAAAATTCACCAACACATTCAGGCTCAACAGAGAAAGAACTG 1764
236 LeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeu 255
1765 -----AATAGCTTTTGAGTCTCAAAAAGAGATATAACTTCGAAAGAGCAGCTT 1818
256 LysGluGluGlyValGlnArg-----AlaArgGluGluGluGlu----- 268
1819 AAGGAGGAGCTGAATGAAACACAGACACACCTTAAAGAAAGAAAGAGGAGATCGCTTCA 1878
269 LysArgLysGluValThrSerHisPheGlnVal-----ThrLeuAsnAspIleGlnLeu 286
1879 AAGCAGAGGAGAAATTCACATTTTCAGGCAAGAGAGAGAGTATCTTCTTCGACGT 1938
287 GlnMetGlu-----Gln 290
1939 CAAAGCAGTCTAGAGCTAGAAATCGTCTGCTTCAAAAGAGAAATGTTACTTGTGCG 1998
291 HisAsn-----GluArgAsnSerLysLeuArgGlnGluLeu 302
1999 CATACTTGGAACAGACCTGTGACGGAGAGTAAACAAAGAGGAGACTCAGAGGAC 2058
303 MetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTyxGluLeuArgGluGluHis 322
2059 TTAGACATGCAATGTTACTGCGACAGCATGAATCCATCAAGAACTGGAGTTTCCGCCAC 2118
323 IleAspLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGln 342
2119 CTCACACTATTTCAGAG-----ATGCGGTGTGAGTTGATC-----AGACTGCAA 2163
343 GlnAlaGlnGluMetLeuLysGluAlaGluGluArgHisGlnArgGluLysAspPheLeu 362
2164 CATCAACTGAGCTTACTAACAGCTGATACATAAGAGAGGAGGACCGGAACTAAGA 2223
363 LeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlnThrHis 382
2224 CGGAAACATGTCATGAAGTTTCAGACAGCTTAAGAGTTTGAAGTCTAAAGAACTCCAA 2283
383 LeuLysGlnGlnLeu-----AlaLeuTyxThrGluLysPheGluGluPheGln 398
2284 ATAAAGAGAGTTTACGATACCTGCAAAATTCAAACACAGACAGTACAAAGCATTAAGG 2343
399 AsnThrLeuSerLysSerGluValPheThrPheLysGlnGlnMetGluLysMet 418
2344 AATCACCTACTGGAG-----ACTACACCAAGAGTGAAGCAACAAAGCTGTT 2388
419 ThrLysLysLysLysLysLeuGluLysGluThrThrMetTyxArgSerArgTrpGluSer 438

2389 CTGAAAGAGACTCAAG-----GAGGAACAGACTCGGAAGATTAGCCATCTTGGCTGAGCAG 2442
439 SerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArg----- 453
2443 TATGATCATAGCATTAATGAATGCTCTCCACACAAGCTCTGGTTTGGATGAAGCAGCAG 2502
453 ----- 453
2503 GAAGCAGATGCCAGGTTTGAAGATCGAGCTACAGCAGGAAGCTGGAGCTTGAATGCA 2562
454 -----AspLysGluLeuGluGlyLeu 460
2563 TATCAGAGCAAAATCAAGATCGAGGCTGAGGCCCAACATGATCGAGAGCTTCGAGAGCTG 2622
461 GlnValLysIle-----GlnArgLeuGluLysLeuCysArg 472
2623 GAACAAAGGCTCTCCCTTCGGAGAGCACTCTTAGAACAGAAAGATTGAAGAGAGATGTTG 2682
473 AlaLeuGlnThrGluArgAsnAsp-----LeuAsnLysArgValGlnAsp 487
2683 GCTTTGCAAGATGAAGCAGACAGAACGAATACGATAGCTGCTCGAGCCGAGCCAGAGAA 2742
488 LeuSerAlaGlyGlyGlnGlySer-----LeuThrAsp 498
2743 ATTGAAGCTTTGACTCTGAAGCATGAGATAGTTTATAGTAACATGGTCTTCTCTAAT 2802
499 SerGlyProGluArgArgProGluGly---ProGlyAlaGlnAlaProSer---SerPro 516
2803 CTCTCCCTCAGGCAATTCAGCCACAGCTACCCAGGAGCTTCTAGCTGGTCTCAATCCT 2862
517 ArgValThrGluAlaPro-----CysTyxProGlyAlaProSerThrGluAlaSerGly 534
2863 ACTGGGTTTCAGAGCTCTACTGGGTCTATCCCATGGTGGCACACCAAGCTTGGGT 2922
535 -----GlnThrGlyProGln 539
2923 CATCCGATGCAAGGCGAGCCCA 2946
RESULT 39
US-03-620-312D-248
; Sequence 248, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yundong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt FL_genes Version 1.0
; SEQ ID NO 248
; LENGTH: 7453

QY 402 SerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLys 421
 DB 5306 GAGTGTCTCAAGCAGCGCTTC-----CGCAAGACCATCTACAG 5344
 QY 422 IleLysLysLysGluLysGluThrThrMetTyrArgSerArgTyrGluSerSerAsnLys 441
 DB 5345 GTGGACACACTGAAGCGCGAGCTAGCAGCGCAGCGCGCGCCAGAGAGTGACAT 5404
 QY 442 Ala-----LeuLeuGluMetAlaGluLysThrValArgAspLysGluLeuGlu 458
 DB 5405 GCACGCCAGCACTGGAGCGGAGCAACAAGAGCTGAAGCCAGCTGCAGGAACCTGAG 5464
 QY 459 Gly-----LeuGlnValLysIleGlnArg 466
 DB 5465 GGTGTGTCTCAAGTCTAAGTTCAGGCCACCATCTCAGCCCTGGAGGCCAGATTGGCAG 5524
 QY 467 LeuGluLysLeuLysArgAlaLeuGlnThrGluArgAsnAspLeuAsn----- 482
 DB 5525 CTGGAGGAGCAGCTTGAAGCAGGAAGCCAGGAGCAGCGCCGCAACAATATTAGTCCGT 5584
 QY 483 -----LysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSer 499
 DB 5585 CCACCTGAGAGAGCTGAAGAAATCTT-CATGAGGTTGAGGATGAGCGTCGACACGC 5643
 QY 500 GlyProGluArgArgProGluGlyProGlyAlaGln----- 511
 DB 5644 GGACCAAGTATAAGAGCAGATGGAGAGCGCCAGCGCTCGGATGAAGCAGCTTAAAGCCCA 5703
 QY 512 AlaProSerSerProArgValThrGluAla-----Pro--- 522
 DB 5704 GCTGAGAGAGCAGAGAGAGAGCGCGCTGCCACCGCATCTCGGCGTAACCTCAGCG 5763
 QY 523 -----CysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyPro 538
 DB 5764 GGAAGTGGATGATGCCACCGAGCGCAAGAGGCGCTGAGCGCGAGGTGAG---CACCCCT 5820
 QY 539 GlnGluProThrSerAla 544
 DB 5821 GAAGAACCGGTGAGCGG 5838

RESULT 41

US-09-220-132-79
 ; Sequence 79, Application US/09220132
 ; Patent No. 6506607
 ; GENERAL INFORMATION:
 ; APPLICANT: Shvian, Andrew W.
 ; TITLE OF INVENTION: OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER
 ; FILE REFERENCE: 07334-074001
 ; CURRENT APPLICATION NUMBER: US/09/220,132
 ; CURRENT FILING DATE: 1998-12-23
 ; PRIOR APPLICATION NUMBER: US 60/079,303
 ; PRIOR FILING DATE: 1998-03-25
 ; PRIOR APPLICATION NUMBER: US 60/068,821
 ; PRIOR FILING DATE: 1997-12-24
 ; NUMBER OF SEQ ID NOS: 191
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 79
 ; LENGTH: 5857
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-220-132-79

Alignment Scores:
 Pred. No.: 9,17e-11 Length: 5857
 Score: 246.00 Matches: 114
 Percent Similarity: 42.31% Conservative: 84
 Best Local Similarity: 24.36% Mismatches: 152
 Query Match: 118
 DB: 20

US-10-023-523-44 (1-546) x US-09-220-132-79 (1-5857)

QY 133 LysGluProSerLysGlyAspProAsnThrGluGluLeuArgGlnSerAspGluValGly 152
 DB 2440 AAAGCCAGTTCGGAAGTAAATCGAATAATGAAGAACTTAGACAGCAGCTTGAGCGAGCT 2499
 QY 153 AspArg-----AspHisArgArgProGlnGlnLysLysLysLysLysLysGly 167
 DB 2500 GAGAAACAGATTAAACATTTAGAGATTGAAAGAAATGCTGAAAGTAGCAGGCTAGTAGC 2559
 QY 168 LeuGlyLysGlu-----IleThrLeuLeuMetGlnThrLeuAsn 180
 DB 2560 ATTACAGAGAGAGCTCCAGGGGAGAGAGCTAAAGCTTACTAACCTTCAGGAAATTTGAGT 2619
 QY 181 ThrLeuSer-----ThrProGluGluLysLeuAlaAlaLeuLysLysLysTyr 196
 DB 2620 GAAGTCACTCAAGTGAAGAGACATTTGGAAGAAAGAACTTCAGATTTTGAAGAAAGTTT 2679
 QY 197 AlaGluLeuLeuGluGlu-----HisArgAsnSerGlnLysGlnMetLysLeu 212
 DB 2680 GCTGAAGCTTCAGAGGAGGAGCTCTCTGTTTCAAGAAATATGCAAGAAACTGTAAATAAG 2739
 QY 213 LeuGlnLysLysGlnSerGlnLeu----- 220
 DB 2740 TTACACCAAGAGGAGGAGACATTTAAACATGCTGCTCTCTGACTTGGAGAGCTGAGAGAA 2799
 QY 221 -----ValGlnGluLysAspHisLeuArgGlyGluHisSer 232
 DB 2800 AACTTAGCAGATATGGAGGCAAAATTTAGAGAGAAAGATGAG-----AGAGAGAG 2850
 QY 233 LysAlaValLeuAlaArgSerLysLeuGluSerLeuLysArgGluLeuGlnArg----- 250
 DB 2851 CAGCTGATAAGGCAAGGAAAGAACTGGAAGAAATGACATTCAGAAATATGAAGATCTCA 2910
 QY 251 -----HisAsnArgSer-----LeuLysGluGluGly 259
 DB 2911 GGAGATAACTCTCTCAGCTGACAAAAAATGAACGATGAATTACCTCTGAAAGAAAGAGAT 2970
 QY 260 ValGlnArgAlaArgGluGluGluLysGluValThrSerHisPheGlnVal 279
 DB 2971 GTAGAGATTACAGCTAAACTTACAAAGGCTAATGAAATGCAAGTTTCTGCAAAAA 3030
 QY 280 ThrLeuAsnAspIleGlnLeuGlnMetGluGln-----HisAsnGluArgAsnSerLysLeu 298
 DB 3031 AGTATTGAGGACATGACTGTCAAGCTGAAAGCTGAAAGAGAGAGAGAGAGAGAGAT 3090
 QY 299 ArgGlnGluAsnMetGluLeuAlaGluLysLysLysLeuIleGluGlnTyrGluLeu 318
 DB 3091 GAGGAGAAAGAAAGAAATTTGGAGAGAAATTTGCGACCTGGAAGAAAGAAATGGAACA 3150
 QY 319 ArgGluGluHisIleAspLysVal----- 326
 DB 3151 AGCCACACACAGTGTCAAGAGCTGAAAGCCAGGATGAGAGAGCCACTTCTGAGACAAA 3210
 QY 327 PheLysHisLysAsp-----LeuGlnGlnGlnLeuValAspAla-----Lys 340
 DB 3211 ACCAAGCATGAAGAAATCTTACAGAACTCCAGAGAGCGTGTGTCACAGAGGACAG 3270
 QY 341 LeuGlnGlnAlaGlnGlu-----MetLeuLysGluAlaGluGlu-----ArgHis 355
 DB 3271 CTGAAGGCGCAGCGGAGGAGAACAGTGGCTTGTGCGAGGAGCTGGAGGAGCTGAGAAAG 3330
 QY 356 GlnArgGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeu 375
 DB 3331 CAGCCGCAAGAGCAAGCTGCTCAACAGCGGAGATGCCATGCGATATGGAACAG 3390
 QY 376 MetLysGlnGln-----GluThrHisLeu 383
 DB 3391 ATGACCAAGAGAGAGAGCTGAGACTCTGCGCTCTTGGAGGAGCAGCAAGCAAAATGCA 3450
 QY 384 LysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThr-----Leu 401
 DB 3451 AAATACAGATGATTTGGACACACTTAAAGAAAGAACTTGAAGAAATGGAAGAGCTG 3510

```
QY 402 SerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLys 421
DB 3511 AACAAATCAAAAGAACTCTGACTAGAGAAATCAAAATGGAAGATTAGGAAGAA 3570
QY 422 IleLysLysLeuGluLysGluThr 430
DB 3571 ATAGAAACCTTAAGCAGGACGACCTCAGAGTCCACAGCTTTTCAGCGTTGCAAGAA 3630
QY 431 -----MetTyxArgSerArgTTPGluSerSerAsnLysAla 442
DB 3631 GAGPACGTTAACTTGCTGAGGAGTGGGAGAGCAGGAGTCAAGATCAACAA 3690
QY 443 LeuLeuGluMetAlaGluGluLysThrValArgAspLysGluLeuGluGluVal 462
DB 3691 AAGCTGGAA-----GAAGAAGATCTGTCTCAATAATCAGTTG-----TTAGAAATG 3738
QY 463 LysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsn 482
DB 3739 AAA-----AAAGAGATCCAGTTTCAATAAGCAGGATGAAGAGAAAGTCTTCTTCAG 3795
QY 483 LysArgValGlnAspLeuSerAla 490
DB 3796 AAATCCATCAGTATAAATAGTGCC 3819

RESULT 42
US-09-252-991A-15560/C
; Sequence 15560, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15560
; LENGTH: 3402
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15560

Alignment Scores:
Pred. No.: 7,47e-11 Length: 3402
Score: 243.00 Matches: 145
Percent Similarity: 40.03% Conservative: 106
Best Local Similarity: 23.13% Mismatches: 218
Query Match: 8.73% Indels: 158
DB: 4 Gaps: 27

US-10-023-523-44 (1-546) x US-09-252-991A-15560 (1-3402)
QY 19 SerProGlyGlnProGluAlaGlyProGluGluValAlaGlnGlu---ArgProSerGlnAla 37
DB 3121 GCGCCTGAAGACGATCAAGTGGC-----GGGGTTCAAGTCTCTTCGATCCGACAC 3068
QY 38 AlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGly 57
DB 3067 GTGTAATCTTCCCGACCAACATGGC---GGCAGTGTGTAGGGCCGCAACCGCTGGCGGAAGTC 3011
QY 58 AlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArg 77
DB 3010 GAACATCATCGACCGGTGGCTGGTGGTGTGGCGGNAAGCTCGGCGAAGAACCTCGCGCG 2951
QY 78 GlnLeuGluAspIleLeuSerThrTyrcysValAspAsnGlnGly---ProGly 96
DB 2950 CGAGTCATGACGATGTATCTTCAACGGCTCGAATACCCGCAAGCCGGTGGAGCAGGC 2891
QY 97 GluAspGlyAlaGln-GlyGluProAlaGlu-----ProGluAs 109

2890 GACATCGAGCTGATCTTCGACACCCGAGACCACTGGTGGGGAATACGCCAGTA 2831
109 pAlaGluLysSer---ArgThrTyrrValAlaArgAsnGlyGluProGluProThProVa 128
2830 CGCGGAGATATCCATTGCGCGGGGTCTCGCGGATGG-----CAGAACACCTATT 2777
128 lValAsnGlyGluLysGluProSerLysGly----- 138
2776 CTTCAACGCGACCAAGTCCGCGCGCGACATCCACGACATCTTCTCGGACCGGCT 2717
139 -AspProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAspHisArgAr 158
2716 GGGCGCGCGCAGCTACTCGATCATCGAACAGGCGCATCTCCAAGCTGATCGAGCGCG 2657
158 gProGlnGluLysLys-----LysAlaLysGlyLeuGlyLysGluLeuThrLe 174
2656 TCCGGAAGACCTGCGCAACTTCATCGAGGAAGCCCGGCGCATTTTCCAAGTACAGAGCG 2597
174 uLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCys-- 193
2596 CCGCGCGGAACCGAAGCGCGCATCCGTCGACCCGAGAAACCTTGGCACCGCTCACCGA 2537
194 -----LysLysTyrrAlaGl 198
2536 CTTGCGCGAAGAGCTGGGCGCGCAACTGGAACCGCTGCACGCGAGGCCAGTCGCGGA 2477
198 uLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLys-----LeuLe 213
2476 AAAATACAGAGACCAAGCGCGAGCGCGCAGCTCAAGCCAGCCAGCTGGCGCGCGTGC 2417
213 uGlnLysLysGlnSerGlnLeuValGlnLysAspHisLeuArgGlyGluHisSerLy 233
2416 CTGCGCGACCTGAACGAGCAGCTCGGTGCGCGCGCGGTCTCATCGCGCAGCAGGAT 2357
233 sAlaValLeuAlaArgSerLysLeuGluSerLysCysArgGluLeuGlnArgHisAsnAr 253
2356 CGCC-----TTGAGGCGCTTGTGCGCGAG---CAGCGCGT----- 2323
253 gSerLeuLysGluGluGlyValGlnArgAlaArg-----GluGluGluGluLy 269
2322 -----GCCGATGCGCGCATCGAACGGCTGCGCGCGGCCCATCACGAGTTGTCCGACG 2270
269 sArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeu-----Gl 287
2269 CTTCAACGAGTGCAGGCACGCTTCTATTGCGTGGCGCGGACATCGCCCGGGTCGAGCA 2210
287 nMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGluLeuAlaGl 307
2209 GAGCATCCAGATCGCGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2150
307 uArgLeuLysLysLeuLeuGluGlnTyr-----GluLe 318
2149 ACGGACCGCGCAGGAAACCGAATCGCACCTCGGCGCATGACCGTACCTTGTCTCGCGACCT 2090
318 uArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAs 338
2089 GGCGGAGAA-----ATGGCCATGCTCGCACCGGACCGAAGCAAGGATTCAGCGC 2045
338 pAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGlu-----GluArgHi 355
2044 GGCGCGCGCGGAAGAGCGGCGCATCGCCCTGGAAACAGCGCGCAGCAGGCGATCGAGCGCTG 1985
355 sGlnArgGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLe 375
1984 GCACGACGATGGATGCCCTTCAACGAGCAGCGCGCAACCCCGCGCGCGCGCGCGCGG 1925
375 uMetLysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrrThrLysPheGl 395
1924 GCACGATCGCGCATCCAGCACCTGAGGACGAGCGCTG----- 1888
395 uGluPheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPhelysGlnLume 415
```


Db 2773 CAA-----TTAGAAATAGAGATTCTCCCTGCAA 2802
Qy' 184 ThrProGluGluLysLeuAlaLeuCysLysTyrAlaGluLeuGluGluHis 203
Db 2803 ACTGTAGAAAGGAGAAACACTGATTACTGAGAACTGCAGCAAACTTTAGAGAAAGTA 2862
Qy 204 ArgAsnSerGlnLysGlnMetLysLeuGluGlnLysGlnSerGlnLeuValGlnGlu 223
Db 2863 AAAACTTTAACTCAAGAAAGATGATCTAAACAACTCCCAAGAAAGCTTGCAAAATGAG 2922
Qy 224 LysAspHisLeuArgGlyGluHisSerLysAlaVal----- 235
Db 2923 AGGGACCACTCAAAAGTGATATTCACGATACCTGTTAACTCATATAGATCAAGAA 2982
Qy 236 LeuAlaArgSerLysLeuGlnSerLysLeuGluGlnArgHisAsn-----Arg 253
Db 2983 CAATTACGAAATGCTCTTCAAGTCTGAAACAACTCAAGAAACAATTAATACACTAAA 3042
Qy 254 SerLeuLysGluGluGlyValGlnArgAlaArgGluGluGluLysArgLysVal 273
Db 3043 TCGAAATTTCTGAGAGATTTCCAGGAATTTGCAATATGGAGAAATACAGGAAACT 3102
Qy 274 ThrSerHisPheGlnValThrLeu-----AsnAspIleGlnLeuGlnMetGluGlnHis 291
Db 3103 AAAGATGAATTTCAAGCAAAAGATGTTGGCATAGATAAAACACAGATTTGGAAGCTAAA 3162
Qy 292 AsnGluArgAsnSerLysLeuGlnGlnGluGlnMetGluLeuAlaGluArgLysLys 311
Db 3163 AATACCAAACTCACTCAAGATGTTAAGGATATGAGATTAATGAGCAACAAAGGAG 3222
Qy 312 -----LeuIleGluGlnThrGluLeuArgGluGluHisIleAspLysValPheLys 328
Db 3223 ATATTTCCTTAATACAGAGAAATAATGAATCCCAACAAATGTTAGAGAGTGTATAGCA 3282
Qy 329 HisLys-----AspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnGln 343
Db 3283 GAAAGGAAACAATTTGAAGACTGACCTAAAGGAAAT---ATTGAAATGACCAATGAAAC 3339
Qy 344 AlaGlnGluMet-----LeuLysGluAlaGluGluArgHisGlnArg 357
Db 3340 CAGGAGAAATTAAGACTCTTGGGATGAACTTTAAAGCAACAGAGATATTTGCACAA 3399
Qy 358 GluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLys 377
Db 3400 GAAAGCAACCATGCCATAAGAAAGAGAGAGCTTCTTAGGACCTGTGCAGACTGGCA 3459
Qy 378 GlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluPhe 397
Db 3460 GAAGTTGAAAGAAACTAAAGGAAAGAGCCAGCAACTCCAGAAAGAAAGCAACACTT 3519
Qy 398 GlnAsnThrLeuSerLysSerGluValPheThrThrPheLysGlnGluMetGluLys 417
Db 3520 CTTAATGTA-----CAAGAGAGATGAGTGAG 3546
Qy 418 MetThrLysLysLysLysLysLeuGlu-----LysGluThr 429
Db 3547 ATGCAGAAAGAAATTAATGAATAGAGAAATTAAGAAATGAATTAAGAAACAAAGAAATG 3606
Qy 430 ThrMetTyrArgSerArgTTPGluSerSerAsnLysAlaLeuLeuGluMetAlaGluGlu 449
Db 3607 ACATTG-----GAACATATGGAACACAGAGAGCTTCAAGTTCAGTTCAGAA 3651
Qy 450 LysThrValArgAspLysGluLeuGluGlyLeuGlnValLysIleGlnArgLeuLys 469
Db 3652 CTTAATGAAATTTATGAGGAAGTGAAATCTATACCAAAAGAAAGAAAGTTCTTAAGGAA 3711
Qy 470 LeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgValGlnAspLeuSer 489
Db 3712 TTACAGAGTCATTGTAACAGAGAGAGACCACTTAGAGGATATATAGAGAAATTTGAA 3771
Qy 490 AlaGlyGly 492
Db 3772 GCTACAGGC 3780

RESULT 49

US-09-620-312D-130
; Sequence 130, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yundong
; APPLICANT: Wang, Dunrui
; APPLICANT: John, Tillinghast
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt FL_genes Version 1.0
; SEQ ID NO 130
; LENGTH: 8503
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (91)..(8082)
US-09-620-312D-130

Alignment Scores:
Pred. No.: 5,06e-10 Length: 8503
Score: 240.00 Matches: 115
Percent Similarity: 41.80% Conservative: 112
Best Local Similarity: 21.18% Mismatches: 224
Query Match: 8.62% Indels: 92
DB: 4 Gaps: 17

US-10-023-523-44 (1-546) x US-09-620-312D-130 (1-8503)
Qy 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGln-----SerAsn 15
Db 2275 CTAATAAAGAAGTTGAAGAAATGAAGCTTTGCGGAAGAAGTCAATTTCTGCTTTCAGAA 2334
Qy 16 ProLysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSer 35
Db 2335 TTGAATTTTACCTTCTGAAGTAGAAAGCGGTGAGGAAGAGATACAGACAAATCTTGA 2394
Qy 36 GlnAlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysPro 55
Db 2395 GAGCTCCATATAATAACATCAGAAAAGATAAATTTTCTGAAGTAGTTTCATAAGGAG 2454
Qy 56 GluGlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeu 75
Db 2455 AGTAGAGTTCAAGGT-----TTACTTTGAAGAAAT 2484
Qy 76 SerArgGlnLeuGluAspIleuSerThrTyrCysValAspAsnGlnGlyPro 95
Db 2485 GGGAAACAAAGATGACCTAGCAACTACA-----CAGTCGAATTTAAAGCACT 2535

Qy	354	ArgHisGlnArgGlnIuysAspPheLeuLeuIuysGluuIaValGluSerGlnArgMetCys	373
		::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	19767	CAGGATCAGCAGCAGCAG-----CAGGATGAACAGGACAGCAGGAGGAGCAG	19720
Qy	374	GluLeuMetIysGlnGlnGlnThrHisIleuIuysGlnGlnLeuAlaIeuTyThrGluIuys	393
		:::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	19719	GACACAGGAGGACGAGCAGGAGGAGGAGGAGCAG-----GAGCAGGAG	19672
Qy	394	PheGluGluPheGlnAsnThrIeuSerIuysSerSerGluValPheThrThrPhelysGln	413
		:::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	19671	TTAGAGCATCAGGACGAGCAGGATTAGAGGACGAGGAGGAGGAGCAGGAGCAG	19612
Qy	414	GluMetClluysMetThrIuysIleIeIuysLeuGluIuysGluThrThrMetTyArg	433
		:::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	19611	GAGTTAGAGGACGAGCAGGAGGAGGATTAGAGGACGAGGAGCAGGAGTTAGAGGACGAGGAG	19552
Qy	434	SerArgTrpGluSerSerAsnIuysAlaLeuLeuGluMetAlaGluGluIuysThrValArg	453
		:::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	19551	CAGGAGTTAGAGCAGGACGAGCAGGAGTTAGAGGACGAGCAGGAGGATTAGAGGAGCAG	19492
Qy	454	AspIysGluLeuGluGlyIeuGlnValIuysIleGlnArgLeuGluIuysLeuCysArgAla	473
		:::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	19491	GAGCAGGAGTTAGAGGACGAGGAGGAGGAGGAGCAAGCAGCAGGAGGTGGAGAGCAGGAG	19432
Qy	474	LeuGlnThrGluArgAsnAspLeuAsnIuys---ArgValGlnAspLeuSerIaGlyGly	492
		:::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	19431	CAGGACGAGGACGAGCAGGAGTAATTAGAGGAGGTGGAGGACCAAGACCA-GGAGCAGGAGGA	19373
Qy	493	GlnGly	494
Db	19372	GCAGGA	19367

QY	2	LysAsnGlnAspLysLysAsnGly	-----AlaAlaLysGlnSer	14	
Db	20733	AGAAGGAGGACGACGAGGACGGGGCGATGGAACAAAACGTGTGACATCCAAAGTTCA	20674		
QY	15	AsnProLysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgPro	34		
Db	20673	CMACAGCAGCAGGACCCACACAGCAGGAG	-----CCACAG	-----CAGCAGGAGCCA	20626
QY	35	SerGlnAlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLys	54		
Db	20625	CAGCAGCAGGAGCCCTG	-----CAGGAGCCACACAG	20593	
QY	55	ProGluGlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGlu	74		
Db	20592	CAGGAGCCACAGCAGCAGGACCCACAGCAGCAGGAGCCCTGCAGGAGCCACACAGCAG	20533		
QY	75	LeuSerArgGlnLeuGluAspIleLeuSerThrTyCysValAspAsnAsnGlnGly	94		
Db	20532	GAGCCACAGCAGCAGGAGCCCTGCAGGAGCCACACAGCAGGAGCCACACAGCAGGAG	20473		
QY	95	ProGlyGluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArg	114		
Db	20472	CCACAGCAGCAGGAGCCACAGCAGCAGGAGCCACAGCAGGAGCCACAGCAGCAGGAG	20413		
QY	115	ThrTyValAlaArgAsnGlyGluProGluProThrProValValAsnGlyGluLysGlu	134		
Db	20412	CCACAGCAGCAGGAGCCACAGCAGGAGCCACAGCAGGAGCCACAGCAGCAGGAG	20353		
QY	135	ProSerLysGlyAspProAsnThrGluGluIleArgGlnSerAsp	-----GluValGlyAsp	153	
Db	20352	CCACAGCAGGAGCCACAGCAGGAGCCACAGCAGGAGCCACAGCAGCAGGAG	20293		
QY	154	ArgAspHisArgArgProGlnGluLysLysLysAlaLysGlyLeuGlyLysGluIleThr	173		
Db	20292	CCACAGCAGGAGCCACAGCAGGAGCCACAGCAGGAGCCACAGCAGGAGCCACAGCAGGAG	20260		
QY	174	LeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCys	193		
Db	20259	-----CAGCAGCGGAGCCACAGCAGGAGCCACAGCAG	-----	20224	
QY	194	LysLysTyAlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeu	213		
Db	20223	CGGAGCCACAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAG	20164		
QY	214	GlnLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLys	233		
Db	20163	CAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAG	20104		
QY	234	AlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGluArgHisAsnArg	253		
Db	20103	CAG	-----GATGAGCAGCAGGATGAGCAGCAGCAGGATGAG	20065	
QY	254	SerLeuLysGluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluVal	273		
Db	20064	CAGCAGCAGGATGAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAG	20005		
QY	274	ThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGlu	293		
Db	20004	CAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAG	19945		
QY	294	ArgAsnSerLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuIle	313		
Db	19944	CAGCAGCAGGATGAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAG	19885		
QY	314	GluGlnTyGluLeuArgGluGluHisIleAspLysValPheLysAspLeuGln	333		
Db	19884	GAGCAG	-----GAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAG	19828	
QY	334	GlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluGlu	353		
Db	19827	CAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAG	19768		

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 8, 2004, 06:37:54 ; Search time 617.041 Seconds

(without alignments)

3759.096 Million cell updates/sec

Title: US-10-023-523-44

Perfect score: 2785

Sequence: 1 MKQDKKNGAAKQSNPKSSP.....APSTEARGTGPOEPTSARA 546

Scoring table:

BLOSUM62

Xgapop 10.0, Xgapext 0.5

Ygapop 10.0, Ygapext 0.5

Fgapop 6.0, Fgapext 7.0

Delop 6.0, Delext 7.0

Searched: 337363 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US10023523/runat.08062004.063748.29931/app.query.fasta_1.1422
-DB=N.Geneseq.29Jan04 -QMT=fastap -SUFFIX=ring -WINMACH=0.1 -LOPECL=0
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=50 -LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=50
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10023523@cgn.1.755@runat.08062004.063748.29931 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGIOG
-DEV TIMEOUT=170 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N.Geneseq.29Jan04.*

1: Geneseqn1980s.*

2: Geneseqn1930s.*

3: Geneseqn2000s.*

4: Geneseqn2001as.*

5: Geneseqn2001bs.*

6: Geneseqn2002s.*

7: Geneseqn2003as.*

8: Geneseqn2003bs.*

9: Geneseqn2003cs.*

10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2785	100.0	1638	5	Aah26501 Human low
2	2785	100.0	2523	4	Abao9008 Human LDL
3	2785	100.0	2523	4	Aak52874 Human pol
4	2785	100.0	5085	7	Acc46153 Human dit
5	2785	100.0	5577	7	Acc46130 Human dit
6	2693	96.7	4697	2	Aav32839 Human low
7	2693	96.7	4697	5	Aah26496 Human low
8	2552.5	91.7	4722	2	Aav32836 Rabbit lo

9	2552.5	91.7	4722	5	AAH26491	Aah26491	Rabbit lo
10	2435.5	87.5	5284	5	AAH26501	AAH26501	DNA encod
11	1754.5	63.0	2188	4	AAH26501	AAH26501	Human pol
12	1727	62.0	1793	7	AB234807	AB234807	Coding se
13	1688	60.6	1793	3	AAA35218	AAA35218	Human ace
14	1688	60.6	1793	3	AAA35217	AAA35217	Human ace
15	1688	60.6	1793	3	AAH21340	AAH21340	Human low
16	1688	60.6	1793	3	AAH21339	AAH21339	Human low
17	1688	60.6	1793	7	AB297034	AB297034	Human nuc
18	1688	60.6	1793	7	AB297033	AB297033	Human nuc
19	1597	57.3	1854	2	AAH26497	AAH26497	Interleuk
20	1286	46.2	2045	6	AAH15042	AAH15042	Human cDN
21	1286	46.2	2045	6	ABX70301	ABX70301	Human lnn
22	1280	46.0	1618	7	ACD13396	ACD13396	Human DNA
23	1275	45.8	4174	7	ACC68992	ACC68992	Human neu
24	1251	44.9	4723	2	AAH26501	AAH26501	Neurite e
25	1048.5	37.6	1191	2	AAH29064	AAH29064	BRCA1 mod
26	1048.5	37.6	1191	2	AAH26497	AAH26497	CDNA 091-
27	925.5	33.2	823	4	AAH22469	AAH22469	Human bre
28	810	29.1	1501	4	AAH188232	AAH188232	Human pol
29	810	29.1	1501	9	ADC32449	ADC32449	Human nov
30	810	29.1	1782	9	ADC30642	ADC30642	Human nov
31	805	28.9	529	2	AAH88391	AAH88391	EST clone
32	747.5	26.8	115766	7	ACD13448	ACD13448	Human DNA
33	740.5	26.6	22255	5	AAH26497	AAH26497	Human low
34	642.5	23.1	1233	5	AAH26497	AAH26497	DNA encod
35	642.5	23.1	1734	5	AAH26497	AAH26497	DNA encod
36	597.5	21.5	1843	4	ABL10155	ABL10155	Drosophil
37	550	19.7	2526	5	AAH26497	AAH26497	DNA encod
38	524	18.8	4307	4	ABL10154	ABL10154	Drosophil
39	486	17.5	937	6	ABN74519	ABN74519	Bovine em
40	464	16.7	334	8	ACH31687	ACH31687	Human bon
41	368	13.2	710	4	AAH07907	AAH07907	Human cDN
42	352	12.6	405	4	AAH35582	AAH35582	Human car
43	352	12.6	405	9	ADA5661	ADA5661	Human car
44	318	11.4	725	6	ABO55358	ABO55358	Human ova
45	313	11.2	981	6	ABN74510	ABN74510	Bovine em
46	291	10.4	7078	4	ABL03681	ABL03681	Drosophil
47	291	10.4	13085	4	ABL03680	ABL03680	Drosophil
48	286.5	10.3	5943	4	AAH22920	AAH22920	DNA encod
49	285.5	10.3	6149	9	ADC26270	ADC26270	Human NOV
50	284	10.2	5574	6	ABK35570	ABK35570	Gene MYH1

ALIGNMENTS

RESULT 1

AAH26501

ID AAH26501 standard; DNA; 1638 BP.

XX AAH26501;

XX AAH26501;

XX 12-NOV-2001 (first entry)

XX Human low density lipoprotein binding protein 3 (LBP-3) DNA.

XX Low density lipoprotein binding protein 3; LBP-3; LDL; human;

XX atherosclerosis; antiarteriosclerotic; gene therapy; diagnosis; vaccine;

XX ds.

XX Homo sapiens.

XX OS

XX WO200164874-A2.

XX 07-SEP-2001.

XX 28-FEB-2001; 2001WO-US006356.

XX 02-MAR-2000; 2000US-00517849.

XX 14-JUL-2000; 2000US-00616289.

XX (BOST-) BOSTON HEART FOUND INC.

XX

PI Lees AM, Lees RS, Law SW, Arjona AA;
 XX WPI: 2001-565505/63.
 DR P-PSDB; AAB82808.
 XX
 PT New isolated low density lipoprotein binding polypeptide for treating,
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.
 XX
 XX Claim 8; Fig 8A; 143pp; English.
 XX
 CC The present sequence is that of the coding region of the human gene (see
 CC also AAB82808) encoding novel human low density lipoprotein binding
 CC protein 2 (LBP-3, see AAB82808). The gene was isolated from a genomic DNA
 CC library using LBP-3 cDNA as probe. The LBP-3 protein predicted from the
 CC present sequence differs from that predicted from a cDNA clone (see
 CC AAB82804) in that it contains an additional 16 amino acids at its N-
 CC terminus (the cDNA is a 5' truncation) and has Asn at amino acid position
 CC 130 rather than Tyr. LBP-3 nucleic acids are among claimed
 CC polynucleotides of the invention that encode novel polypeptides capable
 CC of binding to native and methylated LDL. Also claimed are isolated LBP
 CC polypeptides, and biologically active fragments and analogues of them, as
 CC well as expression vectors, cells and methods of producing the LBPs.
 CC Methods for determining if an animal is at risk for atherosclerosis,
 CC methods for evaluating an agent for use in treating atherosclerosis, and
 CC methods for treating a cell having an abnormality in structure or
 CC metabolism of LBP are claimed. Pharmaceutical compositions comprising an
 CC LBP polypeptide or nucleic acid, and vaccine compositions, are also
 CC claimed
 XX
 SQ Sequence 1638 BP; 490 A; 417 C; 520 G; 211 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,838-153 Length: 1638
 Score: 2785.00 Matches: 546
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 5 Gaps: 0

US-10-023-523-44 (1-546) x AAB26501 (1-1638)

QY 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerSerPro 20
 DB 1 ATGAGAACCCAGACAAAAGAACGGGGCTGCCAAACAATCCAAATCCAAAAGACAGCCCA 60
 QY 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
 DB 61 GGACAAACCGGAGACGAGGACCCGAGGAGGCCAGGAGCGGCCAGCCAGCGGCTCCCTGCA 120
 QY 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
 DB 121 GTAGAGACAGAGGTCCTCGGACGACGACGAGCTCTCTCGAAGCCGGAGGGTGCTCAAGCC 180
 QY 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
 DB 181 AGAACGGGCTCAGTCTGGGGCCCTTCGTGATGTCTCTGAGGAGCTGAGCGGCCAATCGAA 240
 QY 81 AspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyVala 100
 DB 241 GACATCTAGACACATCTGTGTGACACATPACACAGGGGGGGCCCCCGGACGAGATGGGGCA 300
 QY 101 GlnGlyGluProAlaGluProGluAlaGluLysSerArgThrTyrValAlaArgAsn 120
 DB 301 CAGGGTGAGCGGCTGAACCCGAAATGACAGATGACAGAGATCCCGACCTATGTGGCAGGAAT 360
 QY 121 GlyGluProGluProThrProValValAsnGlyGlyGluProSerLysGlyAspPro 140
 DB 361 GGGGAGCCTGAACCAACTCCAGTAGTCAATGAGAGAGAGAACCTCTCCAGGGGGATCCA 420
 QY 141 AsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHisArgArgProGln 160
 DB 421 AACACAGAAGAGATCCGGCAGAGTCAAGAGTCCGAGGTCGGAGACCCGAGACCAATCGAAGCCACAG 480

QY 161 GluLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsn 180
 DB 481 GAGAAGAAAAAGCCCAAGGGTTTGGGTAAGAGAGATCACGCTTGTGATGAGACATTTGAAT 540
 QY 181 ThrLeuSerThrProGluGlyLysLeuAlaLeuCysLysLysTyrAlaGluLeuLeu 200
 DB 541 ACTCTGAGTACCCAGAGGAGAGCTGGCTGCTCTGTGCAAGAGTATGCTGAATGCTG 600
 QY 201 GluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeu 220
 DB 601 GAGGACCCCGGAATTCACAGAGCAGATGAAGCTCTACAGAAAAAGCAGAGCCAGCTG 660
 QY 221 ValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLys 240
 DB 661 GTCAAGAGAGAGCACCACCTGCGCGGTGAGCAGACAGAGCGGCTGCTGCCCGCAGCAG 720
 QY 241 LeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGlyVal 260
 DB 721 CTTGAGAGCTATGCGCTGAGCTGAGCGGCACACCGCTCTCTCAGGAAGAGAGTGTG 780
 QY 261 GlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnValThr 280
 DB 781 CAGCGGCGCGGAGGAGGAGAGAGCGCAAGAGGAGTGCCTCGCACTTCAGAGTGACA 840
 QY 281 LeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGln 300
 DB 841 CTGAATGACATTCAGCTGAGATGGACAGACATGAGCGCACTCCCAAGTGGGCCAA 900
 QY 301 GluAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTyrGluLeuArgGlu 320
 DB 901 GAGAACATCGAGCTGGCTGAGAGGCTCAAGAAGCTGATTGAGCAGTATGAGCTGCGCAG 960
 QY 321 GluHisIleAspLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLys 340
 DB 961 GAGCATATCGACAAAGTCTTCAACACAAAGGACCTACACAGCAGCTGGTGTGATGCCAAG 1020
 QY 341 LeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGlnArgGlyLysAsp 360
 DB 1021 CTCACAGCGCCAGGAGATGCTTAAGGAGGAGAGAGAGCGGCACCGCGGAGAGGAT 1080
 QY 361 PheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlu 380
 DB 1081 TTTCTCTCAAGAGGACGAGTAGAGTCCCAGAGAGATGTGTGAGCTGATGAAGCAGCAAG 1140
 QY 381 ThrHisLysLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThr 400
 DB 1141 ACCACCTGAGCAACAGCTTGCCCTATACACAGAGAAGTTTGAGGAGTTCCAGAACACA 1200
 QY 401 LeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLys 420
 DB 1201 CTTTCCAAAAGCAGCGAGGTATTTCACCATTTCAAGCAGAGATGGAAAAGATGACTAAG 1260
 QY 421 LysIleLysLysLysLysGluThrThrMetTyrArgSerArgTyrGluSerSerAsn 440
 DB 1261 AGATCAAGAGAGCTGGAGAAAGAACCCACCATGTACCCGCTCCCGTGGAGAGCAGCAAC 1320
 QY 441 LysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGluLeuGlyLeu 460
 DB 1321 AAGGCCCTCTTGAGATGGCTGAGGAGAAAAACAGTCCGGGATAAAGAACTGGAGGCGCTG 1380
 QY 461 GlnValLysIleGlnArgGluGlyLysLeuCysArgAlaLeuGlnThrGluArgAsnAsp 480
 DB 1381 CAGGTAAAAAATCAACGGCTGAGAGAGTGTGCCGGGCACTGCACAGAGGCGCAATGAC 1440
 QY 481 LeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSerGly 500
 DB 1441 CTGAACAAGAGGATACAGGACCTGAGTGTGTGGCCAGGGCTCCCTCACTCAGAGTGGC 1500
 QY 501 ProGluArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGlu 520
 DB 1501 CTTGAGAGAGGAGGAGGCGCTGGGCTCAAGCACCCAGCTCCCGGAGGTTCACAGAA 1560
 QY 521 AlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGlu 540

Db 1561 GCCTCTGTACCCAGGACCGAGCAGAGCATCAGGCGACTGGCTCAAGAG 1620
QY 541 ProThrSerAlaArgAla 546
Db 1621 CCACCTCCGCGAGGCGCC 1638
RESULT 2
ID ABA09008
XX ABA09008 standard; cDNA; 2523 BP.
AC ABA09008;
XX
DT 11-JAN-2002 (first entry)
XX
DE Human LDL binding protein homologue-encoding cDNA, SEQ ID NO:784.
XX
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnery; antitumor; ss.
XX
OS Homo sapiens.
XX
PN WO200157188-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US003800.
XX
PR 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PA (HYSE-) HYSEQ INC.
PI Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-457740/49.
DR P-PSDB; ABB11764.
XX
PT Human proteins and DNA encoding sequences useful for preventing, treating
PT or ameliorating a medical condition in a mammalian subject e.g. arthritis
PT and cancer.
XX
PS Claim 1; Page 693-694; 1963pp; English.
XX
CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of

the invention are useful for preventing, treating or ameliorating medical
conditions, e.g., by protein or gene therapy. Such conditions include
cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
proliferative retinopathy, atherosclerosis, coronary heart disease,
arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
vascular growth. Polypeptides involved with tissue regeneration and
repair (or nucleic acids encoding them) may be used to promote wound
healing (e.g., of burns, incisions and ulcers), while those with
immunomodulatory activities may be used in the treatment of viral,
bacterial and fungal infections in addition to immune disorders.
Polypeptides with growth factor activity may be used in cell cultures to
promote cell growth. For example, such polypeptides may be used to
manipulate stem cells in culture to give rise to neuroepithelial cells
that can be used to augment or replace cells damaged by illness,
autoimmune disease or accidental damage. The polypeptides and nucleotides
may also be used in the diagnosis of the above conditions, and in drug
screening techniques. The present sequence represents a cDNA encoding a
novel human polypeptide of the invention
XX
SQ Sequence 2523 BP; 666 A; 626 C; 794 G; 437 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 4.48e-153 Length: 2523
Score: 2785.00 Matches: 546
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0
US-10-023-523-44 (1-546) x ABA09008 (1-2523)
QY 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerSerPro 20
Db 119 ATGAGAACCAAGACAAAGAGACGGGGCTGCCAACATCCAAATCAAAAGACGCCCA 178
QY 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
Db 179 GGACACCGGAAGCAGGACCCGAGGAGCCAGAGCGGCCAGCCAGCGGGCTCCTGCA 238
QY 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
Db 239 GTAGAGCAGAGAGTCCCGGCAGCAGCCAGGCTCTCGGAGCGGAGGGGCTCAAGCC 298
QY 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
Db 299 AGAACGGCTCAGTCTGGGGCCCTTCGTGATGTCTCTGAGAGCTGAGCCGCACTGGAA 358
QY 81 AspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAla 100
Db 359 GACATCTAGGCACATCTGTGTGGACAATAACCGGGGGGGGGGGGGGGGGGGGGGCA 418
QY 101 GlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsn 120
Db 419 CAGGGTGAGCGGGCTGAACCCGGAAGATGCAGAGAGTCCCGGACCTATGTGGCAAGGAAT 478
QY 121 GlyGluProGluProThrProValValAsnGlyGluLysGluProSerLysGlyAspPro 140
Db 479 GCGGAGCCCTGAACCACTCCAGTAGTCAATGGAGAGAGAACCCCTCCAAAGGGGATCCA 538
QY 141 AsnThrGluGluIleArgGlnSerAspGluValGlyVasArgAspHisArgArgProGln 160
Db 539 AACACAGAGAGAGATCCGGCAGAGTGCAGAGTTCGGAGCCGAGACCATTCAGAGCCACAG 598
QY 161 GluLysGlyAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsn 180
Db 599 GAGAGAAAAAAGCCCAAGGGTTTGGGAGAGAGATCACCGTCTGATGCAGACATTGAAT 658
QY 181 ThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysTyrAlaGlnLeuLeu 200
Db 659 ACTCTGAGTACCCAGAGAGAGAGTGGCTGCTGTGCAAGAGATGCTGAATCTGCTG 718
QY 201 GluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeu 220

Db 719 GAGGAGCAGCCGGAATTCACAGACAGATGAGCTCTACAGAAAAGCAGACCCAGCTG 778
 QY 221 ValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLys 240
 Db 779 GTGCAAGAGAGAGGACCACTCGCGGGTGAGCAGCAGCAAGCCGCTCTGCGCCGAGCAAG 838
 QY 241 LeuGluSerLeuLysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGlyVal 260
 Db 839 CTTGAGACCTATCGGTGAGCTGCAGCGGCAACACCGCTCTCTCAGAGAGAGGTGG 898
 QY 261 GlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnValThr 280
 Db 899 CAGCGGGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 958
 QY 281 LeuAsnAspLysGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGln 300
 Db 959 CTGATGACATTCAGCTGAGATGGAACAGCAACATGAGCGCAACTCCAGCTGCGCAA 1018
 QY 301 GluAsnMetGluLeuAlaGluArgLeuLysLysLeuLysLeuLysGlnTyrGlnLeuArgGlu 320
 Db 1019 GAGAACATGAGCTGAGCTGAGGCTCAAGAGCTGATTCAGCAGTATGAGCTGCGCGAG 1078
 QY 321 GluHisIleAspLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLys 340
 Db 1079 GAGCATATCGCAAGGCTCTCAACACAGAGGCTCAACAGCAGCTGGTGGATGCCAAG 1138
 QY 341 LeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluLysLeuArgHisGlnArgLysAsp 360
 Db 1139 CTCAGCAGGCGCCAGGAGATGCTAAAGAGCGCAGAGAGCGCACCGCGGAGAGGAT 1198
 QY 361 PheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlu 380
 Db 1199 TTTCTCTGAAAGGAGGAGTAGATGCCAGAGATGTGAGCTGATGAGGAGCAGCAAG 1258
 QY 381 ThrHisLeuLysGlnGlnLeuAlaLeuLysThrGluLysPheGluLysPheGlnAsnThr 400
 Db 1259 ACCCACCTGAAGCAACAGCTGCGCTATACACAGAGAGTTTGAGGAGTTCCAGAACACA 1318
 QY 401 LeuSerLysSerSerGluValPheThrThrPheLysGluMetCysLysMetThrLys 420
 Db 1319 CTTTCCAAAGCAGCAGGATTCACACACATTCAAGCAGAGATGAAAGATGACTAAG 1378
 QY 421 LysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArgTyrGluSerSerAsn 440
 Db 1379 AAGATCAAGAGCTGGAGAAAGAAACACCATCTACCGGTCCCGTGGGAGAGCAGCAAC 1438
 QY 441 LysAlaLeuLeuGluMetAlaGluLysThrValArgAspLysGluLeuGluGlyLeu 460
 Db 1439 AAGCCCTGCTTGAGTGGCTGAGGAGAAACAGTCCGGATAAAGAACTGGAGGCGCTG 1498
 QY 461 GlnValLysIleGlnArgLeuLysLeuLysCysArgAlaLeuGlnThrGluArgAsnAsp 480
 Db 1499 CAGCTAAATCCAAAGCTGGAGAGCTGTGCGGCGCACTGCAGACAGAGCGCAATGAC 1558
 QY 481 LeuAsnLysArgValGlnAspLeuSerAlaGlyGlnGlySerLeuThrAspSerGly 500
 Db 1559 CTGACAGAGGGTACAGGACCTGAGTGTGTGGCGAGGCTCTCTACTACAGTGGC 1618
 QY 501 ProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGlu 520
 Db 1619 CCTGAGAGGAGGCGCAGAGGGGCTGGGGCTCAAGCACCCAGCTCCCGGCTCAGAA 1678
 QY 521 AlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThrClyProGlnGlu 540
 Db 1679 GCGCTTGTCTACCCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1738
 QY 541 ProThrSerAlaArgAla 546
 Db 1739 CCCACCTCCGCGAGGGCC 1756

RESULT 3
 AAK52874

ID AAK52874 standard; cDNA; 2523 BP.
 AC AAK52874;
 DT 06-NOV-2001 (first entry)
 DE Human polynucleotide SEQ ID NO 2403.
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation; ss.
 OS Homo sapiens.
 PN WO200157190-A2.
 XX 09-AUG-2001.
 XX 05-FEB-2001; 2001WO-US0004098.
 XX 03-FEB-2000; 2000US-00496914.
 XX 27-APR-2000; 2000US-00560875.
 XX 20-JUN-2000; 2000US-00598075.
 XX 19-JUL-2000; 2000US-00620325.
 XX 01-SEP-2000; 2000US-00654936.
 XX 15-SEP-2000; 2000US-00663561.
 XX 20-OCT-2000; 2000US-00693325.
 XX 30-NOV-2000; 2000US-00728422.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Dmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
 XX Ma Y, Zhao QY, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 XX Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX WPI; 2001-476283/51.
 XX P-PSDB; AAM79741.
 XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
 XX in diagnosis and gene therapy.
 XX Claim 1; Page 4687-4688; 6221pp; English.
 XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
 XX encoded polypeptides (AAM78323-AAM80302) that exhibit activity relating to
 XX cytokine, cell proliferation or cell differentiation or which may induce
 XX production of other cytokines in other cell populations. The
 XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
 XX peptide therapy. The polypeptides have various cytokine-like activities,
 XX e.g. stem cell growth factor activity, haematopoiesis regulating
 XX activity, tissue growth factor activity, immunomodulatory activity and
 XX activin/inhibin activity and may be useful in the diagnosis and/or
 XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
 XX inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
 XX (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
 XX sequence listing were missing at the time of publication
 SQ Sequence 2523 BP; 666 A; 626 C; 794 G; 437 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 4,48e-153 Length: 2523
 Score: 2785.00 Matches: 546
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0
 US-10-023-523-44 (1-546) x AAK52874 (1-2523)
 QY 1 MetLysAsnGlnAspLysLysAsnGlyAlaLysGlnSerAsnProLysSerPro 20
 Db 119 ATGAGAACCAAGACAGACAAAAGAGACGGGGCTGCCAACATCCAAATCCAAAGACGAGCA 178

QY 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
 DB 179 GGACAAACCGAAGCAGGACCGAGGAGCCAGGAGGCGCCAGCCAGGCGCTCTGCA 238
 QY 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
 DB 239 GTAGAAGCAGAAAGTCCCGGAGCAGCAGCAGCTCTCTGGAAGCCGAGGGGCTCAAGCC 298
 QY 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
 DB 299 AGAAGGCTCAGTCTGGGGCCCTCGTGATGCTCTGAGGAGCTGAGCGCCCACTGGAA 358
 QY 81 AspIleLeuSerThrTyrCysValAspAsnGlnGlyProGlyGluAspGlyAla 100
 DB 359 GACATACCTGAGCACAATAGTGTGGACAATAACAGGGGGGGCCCGCGAGGATGGGGCA 418
 QY 101 GlnGlyGluProAlaGluProGluAspAlaGlnLysSerArgThrTyrValAlaArgAsn 120
 DB 419 CAGGTGAGCGCTGAACCCGAGATCCAGAGAAGTCCCGGACCTATGTGGCAAGGAA 478
 QY 121 GlyGluProGluProThrProValValAsnGlyGluLysGluProSerLysGlyAspPro 140
 DB 479 GGGGAGCCCTGAACCACTCCAGTAGTCAATGGAGAGAAGAACCTCCAGGGGGATCCA 538
 QY 141 AsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHisArgArgProGln 160
 DB 539 AACACAGAAGAGATCCCGCAGGTGACGAGTCCGAGACCGAGACCATCGAAGGCCACAG 598
 QY 161 GluLysLysLysAlaLysGlyLysGluIleThrLeuLeuMetGlnThrLeuAsn 180
 DB 599 CAGAAGAAAGAACCAAGGGTTGGGAAGCAGATCAGTTGCTGTGATCGACACATTGAAT 658
 QY 181 ThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysLysTyrAlaGluLeuLeu 200
 DB 659 ACTCTGAGTACCCAGAGGAAAGCTGCTCTCTGTGCAAGAAGTATGCTGAATCTGCTG 718
 QY 201 GluGluHisArgAsnSerGlnLysGlnMetLysLeuGlnLysLysGlnSerGlnLeu 220
 DB 719 GAGGAGCACCGGAATTCACAGACGAGATGAAGTCTCTCAAGAAAGAGCAGGCGCAGCTG 778
 QY 221 ValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLys 240
 DB 779 GTGCAAGAGAAGGACCACTCGCGGTGAGCAGACAGCAAGCGCTCTGCGCCGCGCAAG 838
 QY 241 LeuGluSerLeuCysArgGluLeuGluArgHisAsnArgSerLeuLysGluGlyVal 260
 DB 839 CTTGAGACCTATGCGGTGAGCTGAGCGGCACAAACCTCTCTCAAGAAAGAGGTGTG 898
 QY 261 GlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnValThr 280
 DB 899 CAGCGGGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 958
 QY 281 LeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGln 300
 DB 959 CTGAATGACATTCAGTCTGATGGAACAGACCAATGAGCGCAACTCCAAGCTCGCCAA 1018
 QY 301 GluAsnMetGluLeuAlaGluArgLysLysLysLeuIleGluGlnTyrGluLeuArgGlu 320
 DB 1019 GAGAACATGAGCTGGCTGAGGAGCTCAGAGCTGATGATGAGCAGTATGAGCTGCGCGAG 1078
 QY 321 GluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeuValAspAlaLys 340
 DB 1079 GAGCATATCGAACAAGTCTTCAAAACAAGGACCTACAAACAGCAGCTGGTGGATGCCAAG 1138
 QY 341 LeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGlnArgGluLysAsp 360
 DB 1139 CTCACAGAGCCCGAGGAGATGCTAAAGGAGCAGAGAGCGGACCGAGGAGAGGAT 1198
 QY 361 PheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlu 380
 DB 1199 TTTCTCTGAAAGAGGAGTATAGTCCAGAGGATGTGTGAGCTGATGAAGCAGCAAGAG 1258

QY 381 ThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThr 400
 DB 1259 ACCCACTGAAGCAACAGCTTGCCCTATACACAGAGAAGTTTGAGGAGTTCCAGAACACA 1318
 QY 401 LeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLys 420
 DB 1319 CTTTCCAAAGCAGCAGGAGTATTCACCATTCAGCAGAGAGATGGAAGAATGACTAAG 1378
 QY 421 LysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArgTrpGluSerSerAsn 440
 DB 1379 AAGATCAAGAAGCTGGAGAAAGAACCAACCATGTACCGGTCCCGGTGGGAGAGCAACAAC 1438
 QY 441 LysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGluLeuGluGlyLeu 460
 DB 1439 AAGGCCCTCTGAGATGGCTGAGGAGAAACAGTCCGGGATAAAGAACTGGAGGCGCTG 1498
 QY 461 GlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAsp 480
 DB 1499 CAGTAAATCAACGGCTGGAGAGCTGTCCGGGCACTGCAGACACAGCGCATGAC 1558
 QY 481 LeuAsnLysArgValGlnAspLeuSerAlaGlyGlnGlySerLeuThrAspSerGly 500
 DB 1559 CTGAACAAGAGGCTCAGGACCTGAGTGTGTGGCCAGGGCTCCCTCACTGACATGGC 1618
 QY 501 ProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGlu 520
 DB 1619 CTTGAGAGAGGCCAGAGGGGCTTGGGGCTCAAGCACCAAGCTCCCGGAGGTCAAGAA 1678
 QY 521 AlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGlu 540
 DB 1679 GCGCCTTGCTACCCAGGAGCAGCCAGCAGACAGACATCAGCCAGACTGGGCTCAAGAG 1738
 QY 541 ProThrSerAlaArgAla 546
 DB 1739 CCCACCTCCGCCAGGGCC 1756
 RESULT 4
 ID ACC46153 standard; cDNA; 5085 BP.
 XX ACC46153;
 XX 02-JUN-2003 (first entry)
 XX Human dithp extracellular signalling protein-encoding cDNA.
 DE Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;
 KW cancer; cell proliferative disorder; autoimmune disorder;
 KW inflammatory disorder; infection; hormonal disorder; metabolic disorder;
 KW neurological disorder; gastrointestinal disorder; transport disorder;
 KW connective tissue disorder; drug screening; proteome analysis;
 KW gene therapy; antisense therapy; genotyping; transgenic animal; knock in;
 KW disease model; toxicological testing; transcript imaging;
 KW extracellular signalling; gene; ss.
 XX Homo sapiens.
 OS WO200297031-A2.
 PN 05-DEC-2002.
 PD 27-MAR-2002; 2002WO-US010056.
 XX 28-MAR-2001; 2001US-0279619P.
 PR 29-MAR-2001; 2001US-0280067P.
 PR 29-MAR-2001; 2001US-0280068P.
 PR 16-MAY-2001; 2001US-0291280P.
 PR 17-MAY-2001; 2001US-0291829P.
 PR 17-MAY-2001; 2001US-0291849P.
 PR 19-JUN-2001; 2001US-0299428P.
 PR 20-JUN-2001; 2001US-0299776P.
 PR 20-JUN-2001; 2001US-0300001P.
 XX

(INCY-) INCYTE GENOMICS INC.

PA Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
 XX PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Ameshey SR;
 PI Daugherty SC, Dan TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;
 PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
 PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
 XX DR WPI; 2003-129518/12.
 DR P-PSDB; ABR41210.

Novel human diagnostic and therapeutic polypeptide useful for identifying
 PT test compound which specifically binds to a polypeptide encoded by human
 PT diagnostic and therapeutic polynucleotide, and to induce antibodies.

XX Claim 2; SEQ ID NO 74; 591pp; English.

XX The invention relates to novel human diagnostic and therapeutic
 CC polynucleotides designated dithp (ACC46080-ACC46749) and to their encoded
 CC proteins (DITHP; ABR41136-ABR41812). The invention also relates to
 CC polynucleotide sequences at least 90% identical to the dithp cDNA
 CC sequences of the invention; recombinant vectors, host cells and
 CC transgenic organisms comprising a dithp nucleic acid sequence; the
 CC recombinant production of DITHP proteins; antibodies specific for DITHP
 CC proteins; microarrays comprising dithp nucleic acid sequences; methods of
 CC detecting dithp nucleotide and protein sequences; methods of screening
 CC for compounds which specifically bind a DITHP protein; and methods of
 CC assessing the toxicity of test compounds using a dithp hybridisation
 CC probe. Dithp nucleic acid sequences and DITHP proteins may be used in the
 CC diagnosis of a wide variety of conditions including cancer and other cell
 CC proliferative disorders; autoimmune or inflammatory disorders; bacterial,
 CC viral, fungal or parasitic infections; hormonal disorders; metabolic
 CC disorders; neurological disorders; gastrointestinal disorders; transport
 CC disorders; and connective tissue disorders. They may also be used to
 CC screen for modulators of protein activity or gene expression. DITHP
 CC proteins can additionally be used in analysis of the proteome of a tissue
 CC or cell type and to induce antibodies. The dithp nucleic acids are
 CC additionally useful in somatic or germline gene therapy of the disorders
 CC mentioned above, as a source of antisense sequences, as a source of
 CC probes and primers, in genotyping and identification of individuals, in
 CC the generation of transgenic animal models of human disease or knock in
 CC humanised animals, in toxicological testing, and in transcript imaging.
 CC The present sequence represents a dithp cDNA encoding a DITHP protein.
 CC which has extracellular signalling activity. Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 5085 BP; 1178 A; 1310 C; 1407 G; 1190 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	9,43e-153	Length:	5085
Score:	2785.00	Matches:	546
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	7	Gaps:	0

US-10-023-523-44 (1-546) x ACC46153 (1-5085)

QY 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerSerPro 20
 DB 280 ATGAGAACCAAGACAAAGAGACGGGCTGCCAACAATCCATCCAAAAGACAGCCCA 339
 QY 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
 DB 340 CGCAACCGGAAGAGAGAGCCCGAGGAGCCAGGAGCGCCAGCCAGCGGCTCTGCA 399
 QY 41 ValGluAlaGluGlyProGlySerGlnAlaProArgLysProGluGlyAlaGlnAla 60
 DB 400 GTAGAGACGAGAGGTCCCGGACGAGCAGAGGCTCTCCGAGCCCGAGGCGGCTCAGCC 459
 QY 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80

DB	460	AGBACGGCTCAGTCTGGGGCCCTTCGTGATGCTCTGAGGAGCTGAGCGCCCACTGGAA	519
QY	81	AspIleLeuSerThrTyrCysValaspAsnGlnGlyGlyProGlyGluaspGlyAla	100
DB	520	GACATACCTAGGACACATACCTGTGTGACATAACACAGGGGGGGCCCGGAGGATGGGCA	579
QY	101	GlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsn	120
DB	580	CAGGCTGACCGGCTGAACCCGAGATGACAGAGAAGTCCCGACCTATGTGCGCAGGAAT	639
QY	121	GlyGluProGluProThrProValValAsnGlyGluLysGluProSerLysGlyAspPro	140
DB	640	GGGAGCCTGAACCAACTCCAGTAGTCAATGAGAGAAGAACCCCTCCAAAGGGGATCCA	699
QY	141	AsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAspHisArgArgProGln	160
DB	700	AACACAGAGAGATCCGGCAGAGTACGAGGTCCGAGAGCCGAGACCCAGAGGCCACAG	759
QY	161	GluLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsn	180
DB	760	GAGACAAAAAGCCAAAGGTTTGGGAAGGAGATCACGTTGCTGATGCAGACATTGAAT	819
QY	181	ThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysTyrAlaGluLeuLeu	200
DB	820	ACTCTGAGTACCCAGAGAGAGAGTGGCTGCTCTGTGCAAGAAGTATCTGAACTGCTG	879
QY	201	GluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeu	220
DB	880	GAGAGACCGGATTCACAGAGCAGATGAAGTCTCTACAGAAAAGCAGAGCCAGCTG	939
QY	221	ValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLys	240
DB	940	GTCAAGAGAGAGACCCACCTGGCGGTGAGACAGCAGAGCGGCTCTGCGCCCGCAGCAG	999
QY	241	LeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGlyVal	260
DB	1000	CTTGAGAGCCTATCCGCTGAGTGCAGCGGCACACCCGCTCCCTCAGGAAGAAGTGTG	1059
QY	261	GlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnValThr	280
DB	1060	CAGCGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1119
QY	281	LeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGln	300
DB	1120	CTGAATGACATTGAGTGCAGTGGAGAACAGCAGCAATGAGCGCAATCCCAAGGTGCCCA	1179
QY	301	GluAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTyrGluLeuArgGlu	320
DB	1180	GAGAACATGGAGCTGGCTGAGAGGCTCAAGAAGTGTATGAGCAGTATGAGCTGGCGCAG	1239
QY	321	GluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeuValAspAlaLys	340
DB	1240	GAGCATATCGCAAGTCTTCAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1299
QY	341	LeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluGluArgHisGlnArgGluLysAsp	360
DB	1300	CTCCAGCGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1359
QY	361	PheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlu	380
DB	1360	TTTCTCTGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1419
QY	381	ThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThr	400
DB	1420	ACCACCTGAGACAGCAGCTTGCCTATACACAGAGAAGTTTGGAGAGTTCCAGAACACA	1479
QY	401	LeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLys	420
DB	1480	CTTTCCAAAAGCAGCGAGGATTTCCACCATTCAGAGGAGATGGAAAAGATGACTAAG	1539
QY	421	LysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArgTyrGlnSerSerAsn	440

Db 1540 AAGATCAAGAGCTGGAGAAAGAAACACCATGTATCCGGTCCCGTGGGAGAGCAGCAAC 1599
Qy 441 LysAlaLeuLeuMetAlaGluGlyThrValArgAspLysGluLeuGluGlyLeu 460
Db 1600 AAGGCCCTGCTTACAGTGGCTGAGGAGAAACAGTCCGGGATAAAGAACTGGAGGCGCTG 1659
Qy 461 GlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAsp 480
Db 1660 CAGGTAAATAAATCCACCGGCTGGAGAGCTGTGCGGGCACTGCAGACAGAGCGCAATGAC 1719
Qy 481 LeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSerGly 500
Db 1720 CTGAACAAGAGGGTACAGCACTGAGTGTGGTGGCCAGAGGCTCCCTCACTGACAGTGGC 1779
Qy 501 ProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGlu 520
Db 1780 COTGAGAGAGGCCACAGAGGGCGCTGGGGCTCAAGCACCCAGCTCCCGCAGGGTCAACGAA 1839
Qy 521 AlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGlu 540
Db 1840 CGCCTTCTCTACCCAGAGGACCGAGCACAGACATCAGCCAGACTGGGCTCAAGAG 1899
Qy 541 ProThrSerAlaArgAla 546
Db 1900 CCCACCTCCGCGAGGGCC 1917
RESULT 5
ACC46130
ID ACC46130 standard; cDNA; 6577 BP.
AC ACC46130;
XX
DT 02-JUN-2003 (first entry)
DE Human dithep extracellular signalling protein-encoding cDNA.
KW Human; dithep; diagnostic and therapeutic polynucleotide; diagnosis;
KW cancer; cell proliferative disorder; autoimmune disorder;
KW inflammatory disorder; infection; hormonal disorder; metabolic disorder;
KW neurological disorder; gastrointestinal disorder; transport disorder;
KW connective tissue disorder; drug screening; proteome analysis;
KW gene therapy; antisense therapy; genotyping; transgenic animal; knock in;
KW disease model; toxicological testing; transcript imaging;
XX extracellular signalling; gene; ss.
OS Homo sapiens.
XX
XX WO200297031-B2.
XX
XX PD 05-DEC-2002.
XX
XX PF 27-MAR-2002; 2002WO-US010056.
XX
XX PR 28-MAR-2001; 2001US-0279619P.
XX PR 29-MAR-2001; 2001US-0280067P.
XX PR 29-MAR-2001; 2001US-0280068P.
XX PR 16-MAY-2001; 2001US-0291280P.
XX PR 17-MAY-2001; 2001US-0291829P.
XX PR 17-MAY-2001; 2001US-0291849P.
XX PR 19-JUN-2001; 2001US-0299428P.
XX PR 20-JUN-2001; 2001US-0299766P.
XX PR 20-JUN-2001; 2001US-0300001P.
XX
XX PA (INCY-1) INCYTE GENOMICS INC.
XX
XX PI Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
XX PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshay SR;
XX PI Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleeefeld Y, Gerstein EH;
XX PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
XX PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
XX
XX DR WPI; 2003-129518/12.
XX DR P-PSDB; ABR41186.

XX Novel human diagnostic and therapeutic polypeptide useful for identifying
PT test compound which specifically binds to a polypeptide encoded by human
PT diagnostic and therapeutic polynucleotide, and to induce antibodies.
XX
PS Claim 2; SEQ ID NO 51; 591pp; English.
XX
CC The invention relates to novel human diagnostic and therapeutic
CC polynucleotides designated dithep (ACC46080-ACC46749) and to their encoded
CC proteins (DITHP; ABR41136-ABR41182). The invention also relates to
CC polynucleotide sequences at least 90% identical to the dithep cDNA
CC sequences of the invention; recombinant vectors, host cells and
CC transgenic organisms comprising a dithep nucleic acid sequence; the
CC recombinant production of DITHP proteins; antibodies specific for DITHP
CC proteins; microarrays comprising dithep nucleic acid sequences; methods of
CC detecting dithep nucleotides and protein sequences; methods of screening
CC for compounds which specifically bind a DITHP protein; and methods of
CC assessing the toxicity of test compounds using a dithep hybridisation
CC probe. Dithep nucleic acid sequences and DITHP proteins may be used in the
CC diagnosis of a wide variety of conditions including cancer and other cell
CC proliferative disorders; autoimmune or inflammatory disorders; bacterial,
CC viral, fungal or parasitic infections; hormonal disorders; metabolic
CC disorders; neurological disorders; gastrointestinal disorders; transport
CC disorders; and connective tissue disorders. They may also be used to
CC screen for modulators of protein activity or gene expression. DITHP
CC proteins can additionally be used in analysis of the proteome of a tissue
CC or cell type and to induce antibodies. The dithep nucleic acids are
CC additionally useful in somatic or germline gene therapy of the disorders
CC mentioned above, as a source of antisense sequences, as a source of
CC probes and primers, in genotyping and identification of individuals, in
CC the generation of transgenic animal models of human disease or knock in
CC humanised animals, in toxicological testing, and in transcript imaging.
CC The present sequence represents a dithep cDNA encoding a DITHP protein
CC which has extracellular signalling activity. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 6577 BP; 1496 A; 1701 C; 1836 G; 1540 T; 0 U; 4 Other;
Alignment Scores:
Pred. No.: 1,24e-152 Length: 6577
Score: 2795.00 Matches: 546
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 7 Gaps: 0
US-10-023-523-44 (1-546) x ACC46130 (1-6577)
Qy 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerPro 20
Db 280 ATGAAGAACCAAGACAAAAAGAACGGGGCTGCCAAACAATCCAATCCAAAAAGCAGCCCA 339
Qy 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
Db 340 GGACAAACCGGAAGCAGGAGCCCGAGGAGCCCGAGGAGCCCGAGGAGGAGGAGGAGGAGG 399
Qy 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
Db 400 GTAGAGACAGAGGTCCCGGAGCAGCAGCGCTCTCTGAGAGCGGAGGAGGAGGAGGAGG 459
Qy 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
Db 460 AGAACCGGCTCAGTCTGGGGCCCTCTCGTGTATGTCTCTGAGGAGCTGAGCGCCCACTGGA 519
Qy 81 AspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyProGlyGluAspGlyAla 100
Db 520 GACATCTAGACACATCTGTGTGACATATACCGGGGGGGGGGGGGGGGGGGGGGGGGGG 579
Qy 101 GlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsn 120
Db 580 CAGGGTGAAGCGGCTGAACCCGAGAGATGCAGAGAGTCCCGAGCCTATGTGTCAGAGAT 639

QY 121 GlyGluProGluProThrProValValIleGluLysGluProSerLysGlyAspPro 140
 Db 640 GGGGAGCCTGAACCACTCCAGTAGCTAATGGAGAGAGAACCTCCCAAGGGGGATCCA 699
 QY 141 AsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHisArgArgProGln 160
 Db 700 AACACAGAGAGATCCGACAGAGTGCAGAGTCGAGACCCAGACCATCGAAGGCCACAG 759
 QY 161 GluLysLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsn 180
 Db 760 GAGAGMAAAACCAAGGTTGGGAGAGAGATCAGCTTGCTGATCGACACATTCGAT 819
 QY 181 ThrLeuSerThrProGluLysLysLeuAlaLeuCysLysLysLysLysLysLysLys 200
 Db 820 ACTCTGAGTACCCAGAGAGAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 879
 QY 201 GluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysLysLysLys 220
 Db 880 GAGGAGCACCGCAATTCAGAGAGCAGATGAGCTCTTACAGAAAGCAGAGCCAGCTG 939
 QY 221 ValGlnGluLysAspHisLysLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLys 240
 Db 940 GTGCAAGAGAGAGACCATCGCGGTGAGCAGCAGAGCGCGTCTGCGCGCGAGCAAG 999
 QY 241 LeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGluGlyVal 260
 Db 1000 CTTGAGAGCTTATGCGTGAGTGCAGCGGCACACCGCTCCCTCAAGAGAGAGGTGTG 1059
 QY 261 GlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnValThr 280
 Db 1060 CAGCGCGCCCGGAGGAGAGAGAGAGCGCAGAGAGGTGACCTCGCACTTCAGGTGACA 1119
 QY 281 LeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGln 300
 Db 1120 CTGAATGACATTCAGCTGAGATGAAACAGCAATGAGCGCACTCCAGCTGGCCCAA 1179
 QY 301 GluAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTyrGluLeuArgGlu 320
 Db 1180 GAGAACATGAGCTGCTGAGAGGCTCAAGAGCTGATTTGAGCAGTATGAGTGGCGAG 1239
 QY 321 GluHisIleAspLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLys 340
 Db 1240 GAGCATATCGACAAAGTCTTCAACACAGAGACCTACACAGCAGCTGTTGGATGCCAAG 1299
 QY 341 LeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGlnArgGluLysAsp 360
 Db 1300 CTTCCAGCGCCCGAGGATGCTAAGGAGGCGAGAGAGCGGCACCGCGGAGAGGAT 1359
 QY 361 PheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGln 380
 Db 1360 TTTCTCTGAAAGAGCGAGTATTCACCATTCAGAGGAGTGTGAGCTGATGAAGCAGAGAG 1419
 QY 381 ThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThr 400
 Db 1420 ACCACCTGAGCAACAGCTTCCCTATACACAGAGAGTTTGAGAGTTCCAGACACA 1479
 QY 401 LeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLys 420
 Db 1480 CTTTCCAAAGCAGCGAGGTATTCACCATTCAGAGCAGGAGATGGAAAAGATGACTAAG 1539
 QY 421 LysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArgTrpGluSerSerAsn 440
 Db 1540 AGATCAAGAGCTGGAGAAAGAACCCACCATGTCCTCCGCTCCGCTGGAGAGCAGCAAC 1599
 QY 441 LysAlaLeuLeuGluMetAlaGluLysThrValArgAspLysGluLeuGluGlyLeu 460
 Db 1600 AAGGCCCTCTTGAGATGCTCAGGAGAAACAGTCGCGGATAAAGAACTGGAGGCGCTG 1659
 QY 461 GlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAsp 480
 Db 1660 CAGGTAAAAAATCAACGGCTGGAGAGCTGTGCGGGCACTGCACAGAGAGCGCAATGAC 1719

QY 481 LeuAsnLysArgValGlnAspLeuSerAlaGlyGlnGlySerLeuThrAspSerGly 500
 Db 1720 CTGAAACAGAGGTTACAGGACCTGAGTGTGGTGCCAGGGCTCCCTCACTCAGAGTGC 1779
 QY 501 ProGluLysArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGlu 520
 Db 1780 CTTGAGAGAGGCGCAGAGGGGCTGGGGCTCAAGCACCAGCTCCCCAGGGTCAAGAA 1839
 QY 521 AlaproCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGlu 540
 Db 1840 GCSCCTTGCTTACCCAGAGACCCGAGCAGAGCATCAGGCGAGCTGGGCTCAAGAG 1899
 QY 541 ProThrSerAlaArgAla 546
 Db 1900 CCCACCTCCGCGAGGGCC 1917
 RESULT 6
 AAV32839
 ID AAV32839 standard; cDNA; 4697 BP.
 XX AC AAV32839;
 XX DT 09-NOV-1998 (first entry)
 XX DE Human low density lipoprotein binding protein LBP-3 cDNA.
 XX KW Low density lipoprotein binding protein; LDL binding protein 3; LBP-3;
 XX KW receptor; human; atherosclerosis; diagnosis; therapy; vaccine; ss.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 XX CDS 3..1595
 XX FT /*tag= a
 XX FT 159..179
 XX FT /*tag= b
 XX FT /*note= "Claim 23"
 XX WO9823282-A1.
 XX PN 04-JUN-1998.
 XX PD 26-NOV-1997; 97WO-US021857.
 XX PF 27-NOV-1996; 96US-0031930P.
 XX PR 03-JUN-1997; 97US-0048547P.
 XX PA (BOST-) BOSTON HEART FOUND INC.
 XX PI Lees AM, Lees RS, Law SW, Arjona AA;
 XX WPI; 1998-322455/29.
 XX P-PSDB; AAW49042.
 XX PT Nucleic acid encoding low density lipoprotein binding proteins and
 PT related vectors - transformed cells, proteins, and modulators of binding,
 PT useful for treatment and diagnosis of atherosclerosis and for identifying
 PT subjects at risk.
 XX PS Claim 10; Fig 17; 47pp; English.
 XX CC This cDNA clone codes for novel human low density lipoprotein (LDL)
 CC binding protein LBP-3 (see AAW49042). It was isolated by screening human
 CC liver, aorta and fetal brain cDNA libraries with rabbit LBP-3 cDNA. cDNA
 CC clones (see AAV2834-39) and encoded rabbit and human LBPs (see AAW49037-
 CC 42) are claimed. An abnormality in LBP metabolism or structure is
 CC diagnostic of a risk for atherosclerosis. The invention provides: methods
 CC for determining if an animal is at risk for atherosclerosis (e.g. for
 CC prenatal screening); methods for treating atherosclerosis (including gene
 CC therapy) using e.g. LBP polypeptides to bind LDL and thereby prevent
 CC formation of atherosclerotic plaque; and methods for treating a cell
 CC having an abnormality in LBP structure or metabolism. Pharmaceutical and
 CC vaccine compositions are also provided, as well as recombinant vectors

CC and host cells used to produce recombinant LBP

SQ Sequence 4697 BP; 1097 A; 1198 C; 1289 G; 1113 T; 0 U; 0 Other;

Alignment Scores:

Prod. No.: 1 95e-147 Length: 4697
Score: 2693.00 Matches: 529
Percent Similarity: 99.81% Conservative: 0
Best Local Similarity: 99.81% Mismatches: 1
Query Match: 96.70% Indels: 0
DB: 2 Gaps: 0

US-10-023-523-44 (1-546) x AAV32839 (1-4697)

QY 17 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 36
Db 3 AAAAGCAGCCAGGACAAACGGAAGCAGGACCGGAGCCAGGAGCGGCCAGCCAG 62
QY 37 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 56
Db 63 GCCGCTCCTGCACTAGAACGAGAGGTCCCGCAGCAGCCAGGCTCCTCGGAGCGGAG 122
QY 57 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 76
Db 123 GGCGCTCAAGCCAGAACCGCTCAGTCTGGGGCCCTTCGTGATGCTCTCTGAGGAGCTGAGC 182
QY 77 ArgGlnLeuGluAspIleLeuSerThrTyrcysValAspAsnGlnGlyGlyProGly 96
Db 183 CCGCAACTGGAGACATCTGACACATATCTGTGGCAATAAACCGGGGGCCCCCGGC 242
QY 97 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 116
Db 243 GAGGATGGGGCACAGGCTGACCGGCTGAACCGAGATGCAGAGAGTCCCGGACCTAT 302
QY 117 ValAlaArgAsnGlyGluProGluProThrProValValAsnGlyGluLysGluProSer 136
Db 303 GTGGCAAGGAATGGGAGGCTCAACCACTCCAGTAGTCTATCGAGAGAGGAACCTCC 362
QY 137 LysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHis 156
Db 363 AAGGGGGATCCAAACACAGAGAGATCCCGCAGAGTGCAGAGTCCGAGACCGACACCAT 422
QY 157 ArgArgProGlnGluLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMet 176
Db 423 CGAAGCCACAGGAGAGAAAGAAAGCAAGGGTTTGGGAGAGAGATCACGTTGCTGATG 482
QY 177 GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaAlaLeuCysLysLysTyr 196
Db 483 CAGACATTGAATCTGAGTACCCAGAGGAGAGCTGGCTCTGTGCAAGAGATAT 542
QY 197 AlaGluLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLys 216
Db 543 GCTGAACTGCTGGAGGAGCACCGGAATTCACAGAGCAGATGAAGCTGTACAGAAAAG 602
QY 217 GlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeu 236
Db 603 CAGAGCCAGTGTGTGAAGAGAGGACCACTCGCGGTGAGCACAGCAAGCCGCTCTG 662
QY 237 AlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLys 256
Db 663 GCCCGCAGAGAGCTTGAGAGCTTATCCGCTGAGTGCAGCGGCACAAACGCTCCCTCAG 722
QY 257 GluGluGlyValGlnArgAlaArgGluGluGluGluLysArgLysGluValThrSerHis 276
Db 723 GAAGAAGGTGTGAGCGGGCCCGGAGGAGGAGAGAGCGAAGCGAGGTGACCTCGCAC 782
QY 277 PheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSer 296
Db 783 TTCAGGTGACACTGAATGACATTCAGCTGCAGATGGAACAGCACAAATGAGCGCACTCC 842
QY 297 LysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTyr 316
Db 843 AAGCTCGGCCAAGAGAACATGGAGCTGGCTGAGAGGCTCAAGAGAGCTGATTGAGCAGTAT 902

QY 317 GluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnGlnLeu 336
Db 903 GAGCTGCGCAGGAGCATATCGACAAAGTCTTCAACACAGAGGACCTTACACAGCAGCTG 962
QY 337 ValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGln 356
Db 963 GTGGATGCCAAGCTCCAGCAGGCCAGGAGATGCTAAAGAGGCGAGAGAGCGCACAG 1022
QY 357 ArgGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMet 376
Db 1023 CGGAGAGAGGATTTCTCTCTGAAAGAGGAGTCCAGAGGATGCTGTAGCTGATG 1082
QY 377 LysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuThrGluLysPheGluGlu 396
Db 1083 AAGCAGCAAGAGAGACCCACCTCGAACCAACAGCTTCCCTTATACAGAGAGTTTGAGAG 1142
QY 397 PheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGlu 416
Db 1143 TTCAGAACACACTTTCACAAAGCAGCGAGGTATTCACCACTTCAAGCAGGAGATGAA 1202
QY 417 LysMetThrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 436
Db 1203 AAGATGACTAAGAAAGATCAAGAGCTCGAGAAAGAAACCAACCATGTACCCGCTCC 1262
QY 437 GluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGlu 456
Db 1263 GAGAGCAGCAACAGAGGCTCTGTGAGTGGCTGAGAGGAAACAGTCCGGGATAAAGAA 1322
QY 457 LeuGluGlyLeuGlnValLysLysLysLysLysLysLysLysLysLysLysLysLys 476
Db 1323 CTCGAGGCGCTGACAGTAAATAATCCAAACGGCTGGAGAGGTGTGCCGGCAGCTCAGACA 1382
QY 477 GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeu 496
Db 1383 GAGCGCAATGACCTGAACAGAGGTTACAGAGCTGAGTGTGTGTGCCAGGGCTCCCTC 1442
QY 497 ThrAspSerGlyProGluArgArgProGluGluGlyProGlyAlaGlnAlaProSerSerPro 516
Db 1443 ACTCAGAGTGGCCCTGAGAGGAGGCCAGAGGGGCTCGGGCTCAAGCACCCAGCTCC 1502
QY 517 ArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThr 536
Db 1503 AGGCTCAGAGAGCGCTTGTCTACCCAGGAGCAGCAGGAGCAGAGATCAGGCCAGACT 1562
QY 537 GlyProGlnGluProThrSerAlaArgAla 546
Db 1563 GGGCTCAAGAGCCACCTCGCCAGGGCC 1592
RESULT 7
AAH26496
ID AAH26496 standard; cDNA; 4697 BP.
AC AAH26496;
XX
DT 12-NOV-2001 (first entry)
DE Human low density lipoprotein binding protein 3 (LBP-3) cDNA.
KW Low density lipoprotein binding protein 3; LBP-3; LDL; human;
KW atherosclerosis; antiarteriosclerotic; gene therapy; diagnosis; vaccine;
KW ss.
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT CDS 3..1596
FT /*tag=a
FT /partial
XX WO200164874-A2.
FN
XX
PD 07-SEP-2001.

XX 28-FEB-2001; 2001WO-US0006356.
 XX PF
 XX 02-MAR-2000; 2000US-00517849.
 XX PR
 XX 14-JUL-2000; 2000US-00616289.
 XX PR
 XX (BOST-) BOSTON HEART FOUND INC.
 XX PA
 XX PI
 XX Lees AM, Lees RS, Law SW, Arjona AA;
 XX DR
 XX WPI; 2001-565505/63.
 XX DR
 XX P-PSDB; AAB82804.
 XX PT
 XX New isolated low density lipoprotein binding polypeptide for treating,
 XX PT
 XX diagnosing and/or identifying therapeutic agents for atherosclerosis.
 XX FS
 XX Example 4; Fig 17; 143pp; English.
 XX CC
 XX The present sequence is that of a partial cDNA encoding novel human low
 XX density lipoprotein binding protein 3 (LBP-3, see AAB82804). Clones were
 XX CC
 XX isolated from human foetal brain, liver and aorta cDNA libraries using
 XX CC
 XX rabbit LBP-3 cDNA as probe. A full-length cDNA clone is given in
 XX CC
 XX AAB2501, and a genomic DNA sequence is given in AAB26497. LBP-3 nucleic
 XX CC
 XX acids are among claimed polynucleotides of the invention that encode
 XX CC
 XX novel polypeptides capable of binding to native and methylated LDL. Also
 XX CC
 XX claimed are isolated LBP polypeptides, and biologically active fragments
 XX CC
 XX and analogues of them, as well as expression vectors, cells and methods
 XX CC
 XX of producing the Lbps. Methods of determining if an animal is at risk for
 XX CC
 XX atherosclerosis, methods for evaluating an agent for use in treating
 XX CC
 XX atherosclerosis, and methods for treating a cell having an abnormality in
 XX CC
 XX structure or metabolism of LBP are claimed. Pharmaceutical compositions
 XX CC
 XX comprising an LBP polypeptide or nucleic acid, and vaccine compositions,
 XX CC
 XX are also claimed
 XX SQ
 SQ Sequence 4697 BP; 1097 A; 1197 C; 1290 G; 1113 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1,95e-147 Length: 4697
 Score: 2693.00 Matches: 529
 Percent Similarity: 99.81% Conservative: 0
 Best Local Similarity: 99.81% Mismatches: 1
 Query Match: 96.70% Indels: 0
 DB: 5 Gaps: 0

US-10-023-523-44 (1-546) x AAB26496 (1-4697)

QY 17 LysSerSerProGlyGlnProGluAlaGlyProGluGlyValaGlnGluArgProSerGln 36
 DB 3 AAAAGCAGCCAGCAGCAACCCGGAAGCAGGACCCGAGGAGCCAGAGCGGCCAGCCAG 62
 QY 37 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 56
 DB 63 GCGGCTCTGTCAGTAGAAGCAGAGGTCCCGGAGCAGGAGCCAGGCTCTCTGAGGAGCGGAG 122
 QY 57 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 76
 DB 123 GGGGCTCAAGCCAGACGGCTCAGTCTGGGGCTCTGCTGAGGAGGTGAGC 182
 QY 77 ArgGlnLeuGluAspLeuSerThrTyrCysValAspAsnGlnGlyProGly 96
 DB 183 CGCAACTGGAAGACATCTAGACATATCTGTGGACATAACCCAGGGGGGCCCGGC 242
 QY 97 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 116
 DB 243 GAGATGGGGCCAGGGTAGCCGGCTGAACCCGAGATGACAGAGAGTCCCGGACCTAT 302
 QY 117 ValAlaArgAsnGlyGluProGluProThrProValValAsnGlyGluLysGluProSer 136
 DB 303 GTGGCAAGGAATGGGAGGCTGAAACCACTCCAGTAGTCTATGAGAGAGGAAGCACTCC 362
 QY 137 LysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHis 156
 DB 363 AAGGGGATCCAAACACAGAGAGATCCCGGAGAGTACGAGGTTCGAGACCCAGACCAT 422

QY 157 ArgArgProGlnGluLysLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMet 176
 DB 423 CCAAGGCCACAGGAGAGAAAAAGCCAGGGTTTGGGAGGAGATCAGTTGCTGATG 482
 QY 177 GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysTyr 196
 DB 483 CAGACATTGAATCTCTGAGTACCCAGAGAGAGCTGGCTGCTGTGTCAAGAGATAT 542
 QY 197 AlagluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLys 216
 DB 543 GCTGAATGCTGGAGGAGACCCGGAATTCACAGAGCAGATGAAGCTTCTACAGAAAAAG 602
 QY 217 GlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeu 236
 DB 603 CAGAGCCAGCTGGTGCAGAGAGAGACCACTCGCGGTGAGCAGCAGCAAGGCGCTCTG 662
 QY 237 AlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLys 256
 DB 663 CCGCCGACAGCTTGAAGAGCTATGCCGTGAGTGCAGCGGCACACCCGCTCCCTCAAG 722
 QY 257 GluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHis 276
 DB 723 GAAGAAGGTGTGACCGGGCCCGGAGAGAGAGAGCGCAGGAGGTGACCTCGCAC 782
 QY 277 PheGlnValThrLeuAsnAspIleGlnLeuMetGluGlnHisAsnGluArgAsnSer 296
 DB 783 TTCCAGGTGACACTGAATGACATTCAGTGCAGATGGAACAGCAGCAATGAGCGCACTCC 842
 QY 297 LysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTyr 316
 DB 843 AAGTGCGCCAAGAGAACATGAGCTGGCTGAGAGGCTCAAGAAGCTGATGAGCAGTAT 902
 QY 317 GluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeu 336
 DB 903 GAGCTGCGCGAGGAGCATATCGCAAAAGTCTTCAACACAGAGACCTTACACAGCAGCTG 962
 QY 337 ValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGln 356
 DB 963 GTGGATGCCAAGCTCCAGCAGGCGCCAGGAGATGCTAAAGAGGCGAAGAGCGGCCAG 1022
 QY 357 ArgGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMet 376
 DB 1023 CGGAGAGAGNTTTCTCTGAAGAGCGCAGTAGAGTCCCGAGAGGATGTGTGAGCTGATG 1082
 QY 377 LysGlnGlnThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGlu 396
 DB 1083 AAGCAGCAAGAGAGCCACCTGAAGCAACAGCTTGGCTTATACACAGAGAGTTTGAGGAG 1142
 QY 397 PheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGlu 416
 DB 1143 TTCCAGAACACACTTTCAAAAGCAGCAGGATATTCACACATTCAGAGCAGGATGGA 1202
 QY 417 LysMetThrLysLysLysLysLeuGluLysGluThrThrMetTyrArgSerArgTyr 436
 DB 1203 AAGTAGCTAAGAGACTCAAGAGCTGGAAGAAACCAACCATGTGTACCGGCTCCCGTGG 1262
 QY 437 GluSerSerLysAlaLeuLeuGluMetAlaGluLysThrValArgAspLysGlu 456
 DB 1263 GAGAGCAGCAACAGGCGCTGCTTGAATGGCTGAGGAGAAACAGTCCCGGATTAAGAA 1322
 QY 457 LeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThr 476
 DB 1323 CTGAGGCGCTGCGAGGTAAATAATCCACCGCTGGAAGAGCTGTGCGGGCAGTGCAGACA 1382
 QY 477 GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlnGlySerLeu 496
 DB 1383 GAGCGCAATGACCTGAACAGAGGGGTACAGAGCTGAGTGTGTGGCCAGGGCTCCCTC 1442
 QY 497 ThrAspSerGlyProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerPro 516
 DB 1443 ACTGACAGTGGCTTGAGAGGAGGCCAGAGGGGCTGGGGCTCAAGCAGCCAGCTCCCCC 1502

QY 517 ArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThr 536
 DB 1503 AGGGTCACAGAGCGCCTTGCTACCCAGGAGCAGCAGCAGAGCATGCGGCGAGCT 1562
 QY 537 GlyProGlnGluProThrSerAlaArgAla 546
 DB 1563 GGGCTCAAGAGCCACCTCCGCGAGGCC 1592

RESULT 8

AAV32836
 ID AAV32836 standard; cDNA; 4722 BP.

AC AAV32836;
 XX DT 09-NOV-1998 (first entry)
 XX Rabbit low density lipoprotein binding protein LBP-3 cDNA.
 DE Low density lipoprotein binding protein; LDL binding protein 2; LBP-2;
 KW receptor; rabbit; atherosclerosis; diagnosis; therapy; vaccine; ss.
 KW Oryctolagus cuniculus.

OS Key Location/Qualifiers
 FH 61..1734
 FT CDS /*tag= a
 FT misc_feature 348..390
 FT /*tag= b
 FT /*note= "Claim 22"

PN WO9823282-A1.

XX 04-JUN-1998.

XX 26-NOV-1997; 97WO-US021857.

XX 27-NOV-1996; 96US-0031930P.

XX 03-JUN-1997; 97US-0048547P.

XX (BOST-) BOSTON HEART FOUND INC.

XX Lees AM, Lees RS, Law SW, Arjona AA;

XX WPI; 1998-322455/28.

XX P-PSDB; AAW49039.

PT Nucleic acid encoding low density lipoprotein binding proteins and related vectors - transformed cells, proteins, and modulators of binding, useful for treatment and diagnosis of atherosclerosis and for identifying subjects at risk.

PS Claim 7; Fig 14; 47pp; English.

CC This cDNA clone codes for novel rabbit low density lipoprotein (LDL) binding protein LBP-3 (see AAW49039). It was isolated by functional screening of a cDNA library, produced from balloon-deendothelialised healing rabbit abdominal aorta mRNA, for clones encoding LBPs able to bind both native and methyl LDL. cDNA clones (see AAV32834-39) and encoded rabbit and human LBPs (see AAW49037-42) are claimed. An abnormality in an aspect of LBP metabolism or structure is diagnostic of a risk for atherosclerosis. The invention provides: methods for determining if an animal is at risk for atherosclerosis (e.g. for prenatal screening); methods for treating atherosclerosis (including gene therapy) using e.g. LBP polypeptides to bind LDL and thereby prevent formation of atherosclerotic plaque; and methods for treating a cell having an abnormality in LBP structure or metabolism. Pharmaceutical and vaccine compositions are also provided, as well as recombinant vectors and host cells used to produce recombinant LBP

XX Sequence 4722 BP; 997 A; 1332 C; 1430 G; 963 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2.91e-139 Length: 4722

Score: 2552.50 Matches: 508
 Percent Similarity: 93.19% Conservative: 12
 Best Local Similarity: 91.04% Mismatches: 25
 Query Match: 91.65% Indels: 13
 DB: 2 Gaps: 3

US-10-023-523-44 (1-546) x AAV32836 (1-4722)

QY 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerSerPro 20
 DB 61 ATGAAGAATCAAGACAAAGAAAGACGGGGCTGCCAAACAGCCCAACCCAAAGACGCGCC 120
 QY 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnLysArgProSerGlnAlaAlaProAla 40
 DB 121 GGACAGCCGGAGCAGGAGCGGAGAGCCAGGGGGCGCCGCGCGCGCGCGCGCC 180
 QY 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
 DB 181 CGAAGACCCGAAGGT---GCCAGCAGCCAGCTCCCGGAGGCGGAGGGGGCTCAAGCC 237
 QY 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
 DB 238 AAAACTGCTAGCCTGGGGGGCTCTGTGTGTCTCTGAGGAGCTGAGCCCGCGCTTGAA 297
 QY 81 AspileLeuSerThrTyrCysValAspAsnAsnGlnGlyProGlyGluAspGlyAla 100
 DB 298 GACATACTCAGTACATCTGTGTGACACAAACACAGGGGGCGCCCGGTGAGGATGGGGTC 357
 QY 101 GlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsn 120
 DB 358 CAGGGTGAGCCCCCTGAACCTGAAGATGCGAGAGTCTCGCGCTATGTGGCAAGAAAT 417
 QY 121 GlyGluProGluPro---ThrProValValAsnGlyGluLysGluProSerLysGlyAsp 139
 DB 418 GGGGAGCGGAGCGCGGCACCCCTAGTCAATGCGAGAAGGAGAGCTCCACAGGCAGAG 477
 QY 140 ProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHisArgArgPro 159
 DB 478 CCGGCGACGGAAGAGATCCGGACAGCGGATGAGTCCGAGACCAGGACCCCGAGGGCCA 537
 QY 160 GlnGluLysLysLysAlaLysGlyLysGluIleThrLeuLeuMetGlnThrLeu 179
 DB 538 CAGGAAAGAGAGAGGCGCAAGGCTCTGGAAAGGAGATCAGCTGTCTGATGACACACTG 597
 QY 180 AsnThrLeuSerThrProGluGluLysLeuAlaLysLysLysLysLysLysLysLys 199
 DB 598 AACACGCTGAGCACCCAGAGAGAGAGCTGCGCGCTCTGTGCAAGAAGTATGCGGAAC 657
 QY 200 LeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGln 219
 DB 658 CTCGAGGAGACCGGAACTCCAGACAGATGAGCTGCTGCGAGAGACAGACAGGCGAG 717
 QY 220 LeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSer 239
 DB 718 CTGGTGACGAGAGAGGACCACTCGCTGGCGAGCACAGCAAGGCGCATCTCTGGCGCGC 777
 QY 240 LysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGluGly 259
 DB 778 AAGCTCGAGAGCCCTGTGCGCGAGCTGCGCGGACCAACCCCTCGCTCAGAGAGAGAGT 837
 QY 260 ValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnVal 279
 DB 838 GTGACAGCGCGCCGAGAGAGGAGAGGAGAGCGCAAGGAGGTGACGTCACTTCCAGATG 897
 QY 280 ThrLeuAsnAspIleGlnLeuGlnMetGlnHisAsnGluArgAsnSerLysLeuArg 299
 DB 898 ACGCTCAACGACATTTCAGCTCAGATGGAGAGACAGACACAGGCGCACTCCAAAGTGGC 957
 QY 300 GlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTyrGluLeuArg 319
 DB 958 CAGGAGACATGGAGCTGCGCGAGCGGCTCAAGAGCTGATTGAGCAGTACGAGCTGCGA 1017
 QY 320 GluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeuValAspAla 339

```

Db 1018 GAAGAGACATCGCAAAAGTCTTCAACACAAAGATCTGACGACGAGCTGGTGGACGCC 1077
Qy 340 LysLeuGlnGlnAlaGlnGluMetLeuLeuGluAlaGluGluHisGlnArgGluLys 359
Db 1078 AAGCTCCAGCAGGCCAGGAGATGCTCAAGAGGCGAGAGCGGCCACCGCGGAGAG 1137
Qy 360 AspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGln 379
Db 1138 GACTTTCTCTGAGGAGGCGCTGGAGTCCAGAGGATGTGCGAGCTGTGAAGCAACAG 1197
Qy 380 GluThrHisLeuLysGlnGlnLeuAlaLeuThrThrGluLysPheGluGluPheGlnAsn 399
Db 1198 GAGACCCACTGAGCAGCAGCTGCTCCCTATACACAGAGAGTTGAGGAGTTCCAGAAC 1257
Qy 400 ThrLeuSerLysSerGluValPheThrThrPheLysGlnGluMetGluLysMetThr 419
Db 1258 ACTCTTTCCAAAAGCAGCAGGAGTCTTCAACACATTCACACAGGAAATGGAAGATGACA 1317
Qy 420 LysLysLysLysLysLeuGluLysGluThrThrMetTyArgSerArgTTPGluSerSer 439
Db 1318 AAGAAGATCAGAGAGCTGGAGAAAGACACCATGTACCGTTCCTCGGTGGAGAGCAGC 1377
Qy 440 AsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGluLeuGluGly 459
Db 1378 AACAAAGCCCTGCTTGAGATGGCTGAGGAGAAACACTCCGGGACAAAGAGCTGGAAGC 1437
Qy 460 LeuGlnValLysLysLeuGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsn 479
Db 1438 CTGAGGTGAATTCACAGCGCTGGAGAGCTGTGCGGGCACTGCAGACAGAGCGCAAT 1497
Qy 480 AspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSer 499
Db 1498 GACTGNACAGAGGGTGCAGGACCTGAGTCCGCTGGCCAGGCGCCCTCTCCGACAGC 1557
Qy 500 GlyProGluArgPro-----GluGlyPro 508
Db 1558 GGTCTGAGCGGAGGCGCAGAGCCGCCACCCCTCCAAAGAGCAGGAGTGTCTCGAGGCGCCC 1617
Qy 509 GlyValaGlnAlaProSerSerProArgValThrGluAlaProCysTyProGlyAlaPro 528
Db 1618 GGGCTCAGTACCCAACTCTCCAGGCGCACAGACGCTCTCTGCTGCCAGGTGCACCC 1677
Qy 529 SerThrGluAlaSerGlnThrGlyProGlnGluProThrSerAlaAlaGala 546
Db 1678 AGCACAGAGCATCAGGCCAGACAGGCGCCCGCAGAGCCACCACTGCCACTGCC 1731

```

RESULT 9

AAH26491 standard; cDNA; 4722 BP.

AAH26491;

12-NOV-2001 (first entry)

Rabbit low density lipoprotein binding protein 3 (LBP-3) cDNA.

Low density lipoprotein binding protein 3; LBP-3; LDL; rabbit; atherosclerosis; antiatherosclerotic; gene therapy; diagnosis; vaccine;

Oryctolagus cuniculus.

Key Location/Qualifiers
CDS 25..1734
/*tag= a

WO200164874-A2.

07-SEP-2001.

28-FEB-2001; 2001WO-US006356.

PR 02-MAR-2000; 2000US-00517849.
PR 14-JUL-2000; 2000US-00616289.
XX (BOST-) BOSTON HEART FOUND INC.
PI Lees AM, Lees RS, Law SW, Arjona AA;
XX WPI; 2001-565505/63.
DR P-PSDB; AAB82801.
XX New isolated low density lipoprotein binding polypeptide for treating,
diagnosing and/or identifying therapeutic agents for atherosclerosis.
XX Claim 5; Fig 14; 143pp; English.

The present sequence is that of cDNA encoding novel rabbit low density lipoprotein binding protein 3 (LBP-3, see AAB82801). The cDNA was isolated following screening of a rabbit cDNA library for clones encoding LBPs that bound to both native low density lipoprotein (LDL) and methyl LDL. The invention provides claimed polynucleotides encoding novel polypeptides which are capable of binding to native and methylated LDL, the isolated polypeptides, termed LBPs, and biologically active fragments and analogues of them, as well as expression vectors, cells and methods of producing the LBPs. Also claimed are methods of determining if an animal is at risk for atherosclerosis, methods for evaluating an agent for use in treating atherosclerosis, and methods for treating a cell having an abnormality in structure or metabolism of LBP. Pharmaceutical compositions comprising an LBP polypeptide or nucleic acid, and vaccine compositions, are also claimed

Sequence 4722 BP; 997 A; 1332 C; 1430 G; 963 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,91e-139 Length: 4722
Score: 2552.50 Matches: 508
Percent Similarity: 93.19% Conservative: 12
Best Local Similarity: 91.04% Mismatches: 25
Query Match: 91.65% Indels: 13
Gaps: 3

US-10-023-523-44 (1-546) x AAH26491 (1-4722)

Qy 1 MethLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerSerPro 20
Db 61 ATGAAGAATCAAGACAAAGAACCGGGCTGCCAAACAGACCCCAACCCCAAGCAGCCCG 120
Qy 21 GlyLynProGluAlaGlyProGluGlyValaGlnGluArgProSerGlnAlaAlaProAla 40
Db 121 GGACAGCCGGAAGCAGAGCGGAGGAGCCAGGGCGGGCCCGCGCGCGCGCGCGCC 180
Qy 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
Db 181 CGAAGCCGGAAGCT---GCCAGCAGCCAGGCTCCCGGAGGCGGAGGGGCTCAAGCC 237
Qy 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluLeuSerArgGlnLeuGlu 80
Db 238 AAAAATCTCTCAGCTGGGGCGCTCTGTGATGTCTCTGAGGAGCTGAGCCCGCAGTTGGAA 297
Qy 81 AspIleLeuSerThrTyrcysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAla 100
Db 298 GACATCTCAGTACATCTGTGGACAAACACAGGGGGCGCGGGTGGAGTGGGGTC 357
Qy 101 GlnGlyLynProAlaGluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsn 120
Db 358 CAGGCTGAGCCCTGAACTGAGATGCAAGATGCAAGAGTCTCGCCCTATGTGGCAAGGAAT 417
Qy 121 GlyLynProGluPro---ThrProValValAsnGlyLysGluProSerLysGlyAsp 139
Db 418 GGGGAGCGGAGCGGCGCACCCCGAGTAGTCAATGGGAGAGGAGACCTCCAGGCGAGAG 477
Qy 140 ProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAspHisArgArgPro 159
Db 478 CCGGGCAGCGAAGAGATCCCGAGCGAGCGATGAGGTGGAGACCGAGACCCAGCGAGGCCA 537

QY 160 GlnGluLysLysLysAlaLysGlyLeuGlyLysGluLeuThrLeuLeuMetGlnThrLeu 179
Db 538 CAGGAAAGAGAGAGGCGCAAGGGTCTCGGAAAGAGATCACTGCTGATGCGAGACATG 597
QY 180 AsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysTyraAlaGluLeu 199
Db 598 AACACGCTGAGCACCAGAGGAGAGCTGGCGCTCTGTGCAAGAGTATGCGAAGCTG 657
QY 200 LeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGln 219
Db 658 CTCGAGGAGCACCAGAACTCGCAGAGCAGATGAAGCTGCTGCAGAGAAAGAGAGCCAG 717
QY 220 LeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSer 239
Db 718 CTGTGTCAGAGAGAGGAGCACCCTCGTGGCGAGCAGAGAGCCATCTCTGGCCCGCAGC 777
QY 240 LysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGly 259
Db 778 AAGCTCGAGAGCTGTGCGGAGCTGCAGCGGCACACCGCTCGCTCAAGGAGAGAGGT 837
QY 260 ValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnVal 279
Db 838 GTGTCAGGAGCCCGAGAGGAGGAGGAGGAGCGCAAGAGGTGAGCTCACACTTCCAGATG 897
QY 280 ThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArg 299
Db 898 ACGTCTCAACGACATTCAGCTGCAGATGGAGCAGCAGCAGCGCACTCCAGCTGCGC 957
QY 300 GlnGluAsnMetGluLeuAlaGluArgLysLysLeuIleGluGlnTyrGluLeuArg 319
Db 958 CAGGAGAACATGGAGCTGGCGAGCGCTCAAGAAGCTGATTTGAGCAGTACGAGCTCGA 1017
QY 320 GluGluHisIleAspLysValPheLysValPheLysHisLysAspLeuGlnGlnLeuValAspAla 339
Db 1018 GAAGAGCACATCGAAGAGTCTTCAACACAGAGATCTGCAGCAGCAGCTGTGGAGCC 1077
QY 340 LysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluGluArgHisGlnArgGluLys 359
Db 1078 AAGCTCAGCAGGCGCCAGGAGATGCTGAAGGAGCAGAGGCGCAGCAGCGGAGAG 1137
QY 360 AspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGln 379
Db 1138 GACTTTCTCTGAGGAGGCGCTGGAGTCCAGAGGATGTGCGAGCTGATGAAGCAACAG 1197
QY 380 GluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsn 399
Db 1198 GAGACCCACCTGAGCAGCAGCTTGCCCTATACACAGAGAGTTTGAGGAGTTCCAGAAC 1257
QY 400 ThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThr 419
Db 1258 ACTCTTTTCCAAAGCAGCGAGGTGTTCCACCACTTCAACAGGAAATGGAAAGATGACA 1317
QY 420 LysLysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArgTyrGluSerSer 439
Db 1318 AAGAAGATCAAGAGCTGGAGAAAGAGACACCATATGATACCTTCCCGTGGGAGAGCAGC 1377
QY 440 AsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGluLeuGly 459
Db 1378 AACAGGCGCTGCTTACATGGCTGAGGAGAAACACATCCGGGACAAAGAGCTGGAAGC 1437
QY 460 LeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsn 479
Db 1438 CTGAGGTGAAATCCAGCGCTGAGAGAGTGTGCGCGGACATGTCAGAGAGAGCGCAAT 1497
QY 480 AspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlnGlnLysLeuThrAspSer 499
Db 1498 GACCTGACACAGAGGTGTCAGGAGCTGAGTCCCGTGGCCAGGCGCCGCTCCGACAGC 1557
QY 500 GlyProGluArgArgPro-----GluGlyPro 508
Db 1558 GGTCTGTGCGGAGGAGCCAGCCCGCCACCTTCAAGGAGCAGGAGGTGTGAGGGCCCC 1617

QY 509 GlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyrProGlyAlaPro 528
Db 1618 GGGGTCTAAGTACCACTCTCCAGGGCCACAGAGCTTCTCTGTCGCGAGTGCACCC 1677
QY 529 SerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAlaArgAla 546
Db 1678 AGCACAGAGCATCAGGCGCAGACAGAGGCGCCAGGAGCCACCTGCACACTGCC 1731
RESULT 10
AAS65026
ID AAS65026 standard; cDNA; 5284 BP.
XX
AC AAS65026;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #830.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
FN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
XX P-PSDB; ABG00839.
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 830; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 5284 BP; 1370 A; 1419 C; 1627 G; 868 T; 0 U; 0 Other;
Alignment Scores: 2.1e-132 Length: 5284
Pred. No.: 2435.50 Matches: 500
Score:

PD 09-AUG-2001.
XX 05-FEB-2001; 2001WO-US004098.
XX 03-FEB-2000; 2000US-00496914.
XX 27-APR-2000; 2000US-00560875.
XX 20-JUN-2000; 2000US-00598075.
XX 19-JUL-2000; 2000US-00620325.
XX 01-SEP-2000; 2000US-00654936.
XX 15-SEP-2000; 2000US-00663561.
XX 20-OCT-2000; 2000US-00693325.
XX 30-NOV-2000; 2000US-00728422.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI xue AU, Yang Y, Wejhrman T, Goodrich R;
XX WPI; 2001-476283/51.
XX P-FSDB; AAM78757.
XX DR
XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
XX in diagnosis and gene therapy.
XX Claim 1; Page 1605-1607; 6221pp; English.
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
XX encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
XX (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
XX sequence listing were missing at the time of publication
XX Sequence 2188 BP; 541 A; 559 C; 669 G; 418 T; 0 U; 1 Other;
SQ
Alignment Scores:
Pred. No.: 3,36e-93 Length: 2188
Score: 1754.50 Matches: 384
Percent Similarity: 78.35% Conservative: 14
Best Local Similarity: 75.59% Mismatches: 60
Query Match: 63.00% Indels: 51
DB: 4 Gaps: 9
US-10-023-523-44 (1-546) x AAK51890 (1-2188)
QY 46 ProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAlaA-gThrAlaGlnSer 65
DB 31 CCGGGTGCACCCACGCGTCCGGTCCAGCAAGTTGGGGAATGCA---ACAGCTTGAAGT 87
QY 66 GlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu---AspIleLeuSer 84
DB 88 GGCTCTAGTGGCCAGAGCTCAGAG-----CTGGAATAGGAATTGGCATCT 132
QY 85 ThrTyrcysValAspAsnAsnGlnGlyClyProGlyGluAspGlyAlaGlnGlyGluPro 104
DB 133 GCTGGCTGTGTG-----GCCCCT 150
QY 105 AlaGluProGluAspAlaGluLysSerArgThrTyrrValAlaAArgAsnGlyGluProGlu 124
DB 151 GCTTCCCTAGTAGTTA-----CAATTCTCTGCTCCCTACGNGGAGCCTTT 198
QY 125 ProThrProValValAsnGlyGluLysGluProSerLysGlyAspProAsnThrGluGlu 144
DB 199 GGGGTATTGTGAGTTCAAGGAGGAGCGTGAAGCACCAGCAGCATCAGCCCATGAG 259

QY 145 IleArgGlnSerAspGluValGlyAspArgAspHisArgArgProGlnGlnLysLysLys 164
DB 259 -----AGT 261
QY 165 AlaLysGlyLeu---GlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSer 183
DB 262 GCTCCTGGCCTGAGAGGTA-AGGCTCAGGCGGAGCTCAGGACCTAGACCTGCTAGT 320
QY 184 ThrProGluGluLysLeuAla-----Ala-LeuCysLysLysLysLys---AlaGluLe 199
DB 321 GATCCCCCACCAGGAAGGCCCCCAAGATGCTCCTGCTCCCTCCTATCCCTGTGCCCA 380
QY 199 uLeuGluGlnHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGl 219
DB 381 GCTGGAGGAGCACCGGAATTCACAGAAGCAGATGAAGCTCTTACAGAAAAAGCAGAGCCA 440
QY 219 nLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSe 239
DB 441 GCTGTGCAAGAGAAGGACCTCGCGCGTGAAGCAGCAAGGCGCTCTCGCCCGGAG 500
QY 239 rLysLeuGluSerLeuCysArgGluLeuGlnHisAsnArgSerLeuLysGluGluGl 259
DB 501 CAAGCTTGAGAGCCTATGCGTGAGCTGCAGCGGCACACCGCTCCCTCAAGGAAGAGG 560
QY 259 yValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnVa 279
DB 561 TGTGCAGCGCGCGGAGGAGGAGAGGCAAGGAGGTGAGCTCGCCTTCAGGT 620
QY 279 lThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluAArgAsnSerLysLeuAr 299
DB 621 GACACTGAATGACATTCAGCTCGAGATGGAACAGACACATGAGCGCACTCCAAAGTGG 680
QY 299 gGlnGluAsnMetGluLeuAlaGluLeuLysLysLysLysLysLysLysLysLysLys 319
DB 681 CCAAGAGAAACATGGAGCTGGCTGAGAGGCTCAAGAAGCTGATTGAGCAGTATGAGTGG 740
QY 319 gGluGluHisIleAspLysValPheLysLysLysLysLysLysLysLysLysLysLys 339
DB 741 CGAGGAGCATATCGACAAAGTCTTCAACACAGGACCTTACACAGCAGCAGCTGGTGGATG 800
QY 339 aLysLeuGlnGlnAlaGlnGluMetLysLysGluAlaGluGluArgHisGlnArgGlu 359
DB 801 CAAGCTCCAGCGCGCCAGAGATGCTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 860
QY 359 sAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGl 379
DB 861 GGAATTTCTCTCTGAAAGAGGAGGAGTCCAGAGGATGCTGTGAGCTGATGAAGCAGCA 920
QY 379 nGluThrHisLeuLysGlnGlnLeuAlaLeuTyrrThrGluLysPheGluGluPheGlnAs 399
DB 921 AGAGACCCACCTGNAACACAGCTTGCCCTATACACAGAGAGAGTTTGAGAGTTCCAGAA 980
QY 399 nThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetTh 419
DB 981 CACACTTCCAAAGCAGGAGGATTTCCACCATTCACAGCAGGAGATGGAAGATGATGAC 1040
QY 419 rLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 439
DB 1041 TAAGAAGATCAGAAAGCTGGAGAGAGAACCACTATGTCCTCCCGTCCCGTGGAGAGCAG 1100
QY 439 rAsnLysAlaLeuLeuGluMetAlaGluLysThrValArgAspLysGluLeuGluGl 459
DB 1101 CAACAAGGCCCTGCTTGAGATGGCTCAGGAGAGAGAGAGTCCGGGATAAAGAACTGGAGGG 1160
QY 459 yLeuGlnValLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 479
DB 1161 CTTGCAAGGTAATAAATCCACCGCTGGAGAGAGTGTGCCCGGACATGTCAGACAGAGCGCAA 1220
QY 479 nAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSe 499
DB 1221 TGACCTGAACAGAGAGGATCAGGACCTGAGTGTGTGTGSCCAGGGCTCCCTCACTGACAG 1280
QY 499 rGlyProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValTh 519

1281 TGGCCCTGAGAGAGCCAGAGGGGCTGGGGCTCAAGCACCCAGCTCCCGGAGTCAAC 1340
 519 rclualaproCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProG1 539
 1341 AGAAGCGGCTTCTACCCAGGAGACCCGAGACAGACAGATCAGGCCAGACTGGGCTCA 1400
 539 nGluProThrSerAlaArgAla 546
 1401 AGAGCCACCTCCGCGAGGGCC 1422
 RESULT 12
 ID ABZ34807/c
 XX ABZ34807 standard; cDNA; 1793 BP.
 AC ABZ34807;
 XX
 DT 04-FEB-2003 (first entry)
 XX
 DE Coding sequence SEQ ID 165, upregulated in osteogenesis.
 XX
 KW Osteopathic; osteogenesis modulator; gene therapy; osteogenesis;
 KW osteoporosis; bone disease; downregulator; human; ss.
 XX
 OS Homo sapiens.
 XX
 FN WO200281745-A2.
 XX
 PD 17-OCT-2002.
 XX
 PF 05-APR-2002; 2002WO-IB002211.
 XX
 XX 05-APR-2001; 2001US-0281400P.
 PR
 XX (AVET) AVENTIS PHARMA SA.
 PA
 PI Garcia T., Roman Roman S., Barton R., Call K., Theilhaber J;
 PI Connolly T., Jackson A., Bushnell SE, Rawadi G;
 XX
 XX WPI; 2003-058567/05.
 DR
 XX Novel isolated nucleic acid upregulated/downregulated in osteogenesis,
 PT useful for bone disease therapy in subject.
 PT
 XX Claim 26; Page 165-166; 237pp; English.
 PS
 XX The present invention relates to novel nucleotide sequences, which are
 CC differentially expressed in models of osteogenesis upon being put in
 CC contact with a stimulator of osteogenesis. The present sequence is one
 CC such sequence. This sequence can be used for diagnosing osteoporosis/bone
 CC disease in a patient, promoting osteogenesis and/or preventing
 CC osteoporosis/bone disease. The present sequence encodes a secreted
 CC protein
 XX
 SQ Sequence 1793 BP; 339 A; 546 C; 444 G; 464 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1,08e-91 Length: 1793
 Score: 1727.00 Matches: 342
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 62.01% Indels: 0
 DB: 7 Gaps: 0
 US-10-023-523-44 (1-546) x ABZ34807 (1-1793)
 QY 205 AsnSerGlnLysGlnMetLysLeuGlnLysGlnSerGlnLeuValGlnGluLys 224
 DB 1791 AATTCACAGACGACATGAAGCTCTCAGAAAAGACGACCGAGCTGGTCCAGAGAAG 1732
 QY 225 AspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeu 244
 DB 1731 GACCACCTGGCGGTGAGCACAGCAAGGCCCTCTCTGGCCCGCAGCAAGCTTGAGAGCCTA 1672

QY 245 CysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArg 264
 DB 1671 TGGCGTGAGCTGCAGCGGCACAAACCGCTCCCTCAAGGAAGAAGGTGTGCAGCGGCGCG 1612
 QY 265 GluGluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIle 284
 DB 1611 GAGGAGGAGGAGGAGCGCAAGGAGGTGACCTCGCACTTCACAGTGCACATGAATGACATT 1552
 QY 285 GlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGlnAsnMetGlu 304
 DB 1551 CAGCTGCAGTGGACACAGCACCAATGAGCGCACTCCAAGCTGCGCCACAGAGACATGGAG 1492
 QY 305 LeuAlaGluArgLeuLysLysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAsp 324
 DB 1491 CTGGCTGAGAGGCTCAAGAAAGCTGATTGAGCAGTATGAGCTCGCGGAGGAGCATATCGAC 1432
 QY 325 LysValPheIleValHisLysAspLeuGlnGlnLeuValAspAlaLysLeuGlnAla 344
 DB 1431 AAGTCTTCAACACACAGGACCTTACACAGCAGCTGGTGGATGCCAGCTCCAGCAGGCC 1372
 QY 345 GlnGluMetLeuLysGluAlaGluArgHisGlnArgGluLysAspPheLeuLeuLys 364
 DB 1371 CAGGAGATGCTAAAGGAGGACAGAGAGCGGCACACAGCGGAGGAAGATTTCTCTGAAA 1312
 QY 365 GluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLysLys 384
 DB 1311 GAGGAGTATGAGTCCAGAGGATGTGTGAGCTGTAGTGAAGCAGCAGAGACCCACCTGAAG 1252
 QY 385 GlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSer 404
 DB 1251 CAACAGCTTCCCTATACACAGAGAGTTTGAGGAGTTCCACAGACACACTTTCCAAAAGC 1192
 QY 405 SerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLys 424
 DB 1191 AGCGAGGTATTCCACCATTCAGCAGGAGATGGAAAAGATCACTAAGAAGATCAAGAAG 1132
 QY 425 LeuGluLysGluThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeu 444
 DB 1131 CTGGAGAAGAAGAACCAACCATGTACCGGTCCCGTGGGAGAGCAGCACAAGGCCCTGCTT 1072
 QY 445 GluMetAlaGluGluLysThrValArgAspLysGluLeuGluGlyLeuGlnValLysIle 464
 DB 1071 GAGATGGCTGAGGAGAAAACAGTCCGGGATAAAGAACTGGAGGGCTTCAGGCTAAAAATC 1012
 QY 465 GlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArg 484
 DB 1011 CAACGCTGGAGAAGCTGTCCGGGCATCTGCAGACAGACGCGCAATGACCTGACACAGAGG 952
 QY 485 ValGlnAspLeuSerAlaGlyGlnGlySerLeuThrAspSerGlyProGluArgArg 504
 DB 951 GTACAGGACCTGAGTGTCTGGTGGCCAGGGCTCCCTCACTGACAGTGGCCCTCAGAGGAGG 892
 QY 505 ProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyr 524
 DB 891 CCAGAGGGGCTGGGGCTCAAGCACCCAGCTCCCGAGGCTCACAGAGCGCTTGTCTAC 832
 QY 525 ProGlyValaProSerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAla 544
 DB 831 CCAGGAGCACCAGGACGACAGACATCAGGCCAGACTGGGCTCAAGAGGCCACCTCCGCG 772
 QY 545 ArgAla 546
 DB 771 AGGGCC 766
 RESULT 13
 AAA35218/c
 ID AAA35218 standard; DNA; 1793 BP.
 XX
 AC AAA35218;
 XX
 DT 28-JUL-2000 (first entry)
 XX

DE Human adenosine receptor related polynucleotide 2nd SEQ ID NO:92..

XX Human; adenosine receptor; low adenosine antisenese oligonucleotide;
 KW phosphorothioate; impaired respiration; inflammation; allergy;
 KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
 KW antiallergic; antiaesthetic; cytostatic; analgesic; impaired airway;
 KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
 KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;
 KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
 KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.

XX Homo sapiens.

XX WO200009525-A2.

XX 24-FEB-2000.

XX 03-AUG-1999; 99WO-US017712.

XX 03-AUG-1998; 98US-0095212P.

XX (UYEC-) UNIV EAST CAROLINA.

XX Nyce JW;

XX WPI; 2000-205971/18.

XX New antisenese oligonucleotides useful for treating e.g. pulmonary
 PT vasoconstriction, inflammation, allergies, asthma, hypertension, or
 PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
 PT cancers.

XX Disclosure; Page 1259-1260; 1343pp; English.

XX The present invention describes a new composition comprising an antisenese
 CC oligonucleotide (ON) with low adenosine (up to 15%), which targets
 CC nucleic acids involved in bronchoconstriction, allergies, and/or
 CC inflammation. The ON can have antiinflammatory, antiallergic,
 CC antiaesthetic, cytostatic and analgesic activities. The compositions are
 CC useful for the treatment of diseases associated with inflammation,
 CC impaired airways, including lung disease and diseases whose secondary
 CC effects afflict the lungs of a subject. They can be used for treating
 CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma,
 CC impeded respiration, respiratory distress syndrome, pain, cystic
 CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive
 CC pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
 CC carcinomas, and cancers which may metastasize to the lungs, including
 CC breast and prostate cancer. The reduction of the adenosine content of the
 CC ONs reduces side effects. The A-containing ONs break down with the
 CC release of deoxyadenosine which activates adenosine receptors causing
 CC bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the
 CC nucleotide sequences given in the sequence listing from the present
 CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last 185
 CC sequences are also called SEQ ID NO:1 to 185, but the sequences differ
 CC from the previously named sequences. SEQ ID NO:11 to 1680 (AAA32323 to
 CC AAA33992) are specifically claimed ONs from the present invention. N.B.
 CC Sequences given in the disclosure of the present invention do not match
 CC up with their corresponding SEQ ID NO: sequences given in the sequence
 CC listing

XX Sequence 1793 BP; 339 A; 546 C; 444 G; 464 T; 0 U; 0 Other;

XX Alignment Scores:

Pred. No.:	2,01e-89	Length:	1793
Score:	1688.00	Matches:	339
Percent Similarity:	99.12%	Conservative:	0
Best Local Similarity:	99.12%	Mismatches:	3
Query Match:	60.61%	Indels:	1
DB:	3	Gaps:	0

US-10-023-523-44 (1-546) x AAA35218 (1-1793)

QY 205 AsnSerGlnLysGlnMetLysLeuGlnLysLysGlnSerGlnLeuValGlnGluLys 224
 AAA35217/c

Db	1791	AATTCAGAGCAGATGAAGCTCTACAGAAAAGCAGAGCCAGCTGGTGCAGAGAAG	1732
Qy	225	AspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeu	244
Db	1731	GACCACCTGGCGGTGAGCAGCAAGCC-GTCTGCGCCGCGCAGCAAGCTTGAGAGCCTA	1673
Qy	245	CysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGlyValGlnArgAlaArg	264
Db	1672	TGCGTGAGGTGACGGGCACACCCGCTCCCTCAAGGAAGAAGGTGTGCAGCGGGCCCGG	1613
Qy	265	GluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIle	284
Db	1612	GAGGAGGAGGAGGAGCGCAAGGAGGTGACCTCGCACTTCCAGGTGACACATGACATT	1553
Qy	285	GlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGlu	304
Db	1552	CAGCTGCAGATGGAACAGCACAATGAGCGCAATCCNAAGCTGGCCNAGAGACATGGAG	1493
Qy	305	LeuAlaGluArgLeuLysLysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAsp	324
Db	1492	CTGCTCAGAGGCTCAAGAAGCTGATTGAGCAGTATGAGCTGGCGGAGGAGCATATCGAC	1433
Qy	325	LysValPheLysHisLysAspLeuGlnGlnLeuValAspAlaLysLeuGlnAla	344
Db	1432	AAAGTCTTCAACACAAAGGACCTACACAGCAGCTGGTGGATGCCAAGCTCCAGCAGGCC	1373
Qy	345	GlnGlnMetLeuLysGluAlaGluArgHisGlnArgGluLysAspPheLeuLysLys	364
Db	1372	CAGAGATGCTAAGGAGGAGCAGAGCGGCACCCAGCGGAGAGGATTTCCTCTGAAA	1313
Qy	365	GluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLeuLys	384
Db	1312	GAGGAGTAGAGTCCAGAGGATGTGTGAGCTGATGAAGCAGCAGAGAGCCACCTGAAG	1253
Qy	385	GlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSer	404
Db	1252	CAACAGCTTGCCCTATACAGAGAGATTGAGAGATTCCAGAACACACTTTCCTCAAAAGC	1193
Qy	405	SerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLys	424
Db	1192	AGCGAGGTATTCCACCACATTCAGCAGGAGATGGAAAAAGATGACTAAGAAGATCAAGAAG	1133
Qy	425	LeuGluLysGluThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeu	444
Db	1132	CTGAGAAAGAAACACCATGTACCGGTCCCGTGGGAGAGCAGCAACAGGCCCTGCTT	1073
Qy	445	GluMetAlaGluGluLysThrValArgAspLysGluLeuGluGlyLeuGlnValLysIle	464
Db	1072	GAGATGCTGAGGAGAAACACAGTCCGGGATAAAGAACTGGAGGGGCTGCAGTAAAAATC	1013
Qy	465	GlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArg	484
Db	1012	CAACGGCTGGAGAAGCTGTGCGGGGCATGCAGACAGAGCGCAATGACCTGAACAAGAGG	953
Qy	485	ValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSerGlyProGluArgArg	504
Db	952	GTACAGGACCTGAGTGTGTGTGGCCAGGGGTCCCTCACTGACAGTGGCCCTGAGAGGAGG	893
Qy	505	ProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyr	524
Db	892	CCAGAGGGGCTCGGGCTCAAGCACCCAGGCTCCCGAGGGGTCAAGAGGCGCTTGTGTAC	833
Qy	525	ProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAla	544
Db	832	CCAGAGACCCGAGCAGAGAGATCAGGCCAGACTGGGCTCAAGAGAGCCACCTCCGCC	773
Qy	545	ArgAla 546	
Db	772	AGGGCC 767	

RESULT 14
 AAA35217/c

Qy 545 ArgAla 546
Db 772 AGGCC 767

RESULT 15

AAF21340/c
ID AAF21340 standard; DNA; 1793 BP.

XX AAF21340;

XX 14-MAR-2001 (first entry)

XX Human low adenosine antisense oligonucleotide related sequence #2907.

KW Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
KW human; airway disorder; bronchoconstriction; lung inflammation;
KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;
KW immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
KW respiratory obstruction; pulmonary obstruction; impeded respiration;
KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;
KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
KW cancer; ss.

XX Homo sapiens.

XX W0200062736-A2.

XX 26-OCT-2000.

XX 24-MAR-2000; 2000WO-US008020.

XX 06-APR-1999; 99US-0127958P.

XX (UYEC-) UNIV EAST CAROLINA.

PA (NYCE/) NYCE J W.

XX Nyce JW;

XX WFI; 2000-679539/66.

XX Low adenosine (A) content antisense oligonucleotides which do not trigger
PT adenosine receptors during metabolism, useful e.g. for treating cancers
PT and respiratory obstructions.

PS Disclosure; Page 1344; 1592pp; English.

XX The present invention describes low adenosine (A) content antisense
CC oligonucleotides and compositions (I) comprising them. In the antisense
CC oligonucleotides the A is replaced by a 'Universal' or alternative base.
CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
CC immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.
CC The antisense oligonucleotides and (I) can be used to down-regulate the
CC expression and/or activity of target polypeptides associated with
CC lung/respiratory disorders and malignancies, such as stimulating and
CC activating peptide factors and transmitters, transcription factors,
CC immunoglobulins and antibodies, antibody receptors, cytokines and
CC chemokines, endogenously produced specific and non-specific enzymes,
CC binding proteins, adhesion molecules and their receptors, cytokine and
CC chemokine receptors, adenosine receptors, bradykinin receptors, central
CC nervous system (CNS) and peripheral nervous and non-nervous system
CC receptors, CNS and peripheral nervous and non-nervous system peptide
CC transmitters, defensins, growth factors, vasoactive peptides and
CC chemokines, binding proteins and malignancy associated proteins. The
CC antisense oligonucleotides may be used in this way to treat disorders
CC including respiratory obstruction (especially pulmonary obstruction
CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or
CC surfactant hypoproduction which are associated with a disease or
CC condition selected from pulmonary vasoconstriction, inflammation,
CC allergies, asthma, impeded respiration, respiratory distress syndrome
CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),

CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
CC fragments and antisense oligonucleotides used in the exemplification of
CC the present invention

XX Sequence 1793 BP; 339 A; 546 C; 444 G; 464 T; 0 U; 0 Other;

Alignment Scores:

Pred. NO.: 2.01e-89 Length: 1793
Score: 1688.00 Matches: 339
Percent Similarity: 99.12% Conservative: 0
Best Local Similarity: 99.12% Mismatches: 3
Query Match: 60.61% Indels: 1
DB: 3 Gaps: 0

US-10-023-523-44 (1-546) x AAF21340 (1-1793)

Qy 205 AsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValcInclulys 224

Db 1791 AATTCACAGAGCAGATCAAGCTCTCTACAGAAAAAGCAGAGCCAGCTGGTCAAGAGAAG 1732

Qy 225 AspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeu 244

Db 1731 GACCACCTCGCGGTGAGCAGCAGCAAGCC-GTCTCTGGCCCGCAGCAAGCTTGAGAGCCTA 1673

Qy 245 CysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGlyValcInArgAlaArg 264

Db 1672 TGCCGTGAGCTGCACGGGCAACAACCGCTCCCTCAAGGAGAGAGGTGTGCAGGGGCCCGG 1613

Qy 265 GluGluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIle 284

Db 1612 GAGGAGGAGGAGAGCGCAAGAGGTGACCTCGCACTTCCAGGTGACACTGAATGACATT 1553

Qy 285 GlnLeuGlnMetGluGlnHisAsnGlnArgAsnSerLysLeuArgGlnGluAsnMetGlu 304

Db 1552 CAGCTGCAGATGGAACAGCACAATGAGCGCAACTCCAAGCTGCGCCAGAGACATGAG 1493

Qy 305 LeuAlaGluArgLeuLysLysLeuIleGluIntyrGluLeuArgGluGluHisIleAsp 324

Db 1492 CTGGCTGAGAGGCTCAAGAAGCTGATTGAGCAGTATGAGCTGCGCAGGAGCATATCGAC 1433

Qy 325 LysValPheLysHisLysAspLeuGlnGlnLeuValAspAlaLysLeuGlnAla 344

Db 1432 AAAAGTCTTCAACACAGAGAGCCTTCAACAGCAGCTGGTGGATGCCAAGCTCCAGAGGCC 1373

Qy 345 GlnGluMetLeuLysGluAlaGluArgHisGlnArgGluLysAspPheLeuLeuLys 364

Db 1372 CAGGAGATGCTAAAGAGGCGCAGAGCGGCACCGCAGCGGAGAGGATTTTCTCTGAAA 1313

Qy 365 GluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnThrHisLeuLys 384

Db 1312 GAGGAGTAGAGTCCAGAGGATGTGTGAGCTGATGAAGCAGCAGAGACCCACCTGAAG 1253

Qy 385 GlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSer 404

Db 1252 CAACAGCTTGCCCTATACAGAGAAGTTTGAGGAGTTCCAGAACACACTTTCACAAAGC 1193

Qy 405 SerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLys 424

Db 1192 ACCGAGGTATTCCACCATTCAGCAGGAGATGGAAGATGACATCAAGAGATCAAGAAG 1133

Qy 425 LeuGluLysGluThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeu 444

Db 1132 CTGGAGAAAGAAACACCATGTACCGGTCCCGGTGGGAGAGCAGCAACAGGCCCTGCTT 1073

Qy 445 GluMetAlaGluGluLysThrValArgAspLysGluLeuGluGlyLeuGlnValLysIle 464

Db 1072 GAGATGGCTGAGAGAGAAACACTCCGGGATAAAGACTGGAGGGCCCTCAGGTAATAATC 1013

Qy 465 GlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArg 484

Db 1012 CAACGGCTGGAGAGCTGTGCCGGGCACTGCAGACAGAGCGCAATGACCTGAACAGAGG 953

QY 485 ValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSerGlyProGluAacArg 504
 DB 952 GTACAGGACCTGAGTGTGGCGCAGGCTCCCTCACTGACAGTGGCCCTTGAGAGGAGG 893
 QY 505 ProGluGlyProGlyAlaGlnAlaProSerProArgValThrGluAlaProCysTyr 524
 DB 892 CCAGAGGGGCTGGGGCTCAAGCACCCAGCTCCCGGAGTCAAGAGCGCTTGCTAC 893
 QY 525 ProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAla 544
 DB 832 CCAGAGACCCGACGACAGAGACATCAGGCCAGACTGGGCTCAAGAGCCACCTCGGCC 773
 QY 545 ArgAla 546
 DB 772 AGGGCC 767

RESULT 16

AAF21339/c

ID AAF21339 standard; DNA; 1793 BP.

XX AC AAF21339;

XX DT 14-MAR-2001 (first entry)

XX DE Human low adenosine antisense oligonucleotide related sequence #2906.

XX KW Low adenosine antisense oligonucleotide; phosphorothioate; allergy;

XX KW human; airway disorder; bronchoconstriction; lung inflammation;

XX KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;

XX KW immunosuppressive; antialasthmatic; analgesic; hypotensive; cytostatic;

XX KW respiratory obstruction; pulmonary obstruction; impeded respiration;

XX KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;

XX KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;

XX KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;

XX KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;

XX KW cancer; ss.

XX OS Homo sapiens.

XX PN WO200062736-A2.

XX PD 26-OCT-2000.

XX PF 24-MAR-2000; 2000WO-US008020.

XX PR 06-APR-1999; 99US-0127958P.

XX PA (UYEC-) UNIV EAST CAROLINA.

XX PA (NYCE/) NYCE J W.

XX FI Nyce JW;

XX DR WPI; 2000-679539/66.

XX PT Low adenosine (A) content antisense oligonucleotides which do not trigger

XX PT adenosine receptors during metabolism, useful e.g. for treating cancers

XX PT and respiratory obstructions.

XX PS Disclosure; Page 1343; 1592pp; English.

XX SS

CC The present invention describes low adenosine (A) content antisense

CC oligonucleotides and compositions (I) comprising them. In the antisense

CC oligonucleotides the A is replaced by a 'universal' or alternative base.

CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,

CC immunosuppressive, antialasthmatic, hypotensive and cytostatic activities.

CC The antisense oligonucleotides and (I) can be used to down-regulate the

CC expression and or activity of target polypeptides associated with

CC lung/respiratory disorders and malignancies, such as stimulating and

CC activating peptide factors and transmitters, transcription factors,

CC immunoglobulins and antibodies, antibody receptors, cytokines and

CC chemokines, endogenously produced specific and non-specific enzymes,

CC binding proteins, adhesion molecules and their receptors, cytokine and

CC chemokine receptors, adenosine receptors, bradykinin receptors, central

CC nervous system (CNS) and peripheral nervous and non-nervous system
 CC receptors, CNS and peripheral nervous and non-nervous system peptide
 CC transmitters, defensins, growth factors, vasoactive peptides and
 CC receptors, binding proteins and malignancy associated proteins. The
 CC antisense oligonucleotides may be used in this way to treat disorders
 CC including respiratory obstruction (especially pulmonary obstruction
 CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or
 CC surfactant hypoproduction which are associated with a disease or
 CC condition selected from pulmonary vasoconstriction, inflammation,
 CC allergies, asthma, impeded respiration, respiratory distress syndrome
 CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
 CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
 CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
 CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
 CC fragments and antisense oligonucleotides used in the exemplification of
 CC the present invention

XX SQ Sequence 1793 BP; 339 A; 546 C; 444 G; 464 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,01e-89 Length: 1793
 Score: 1688.00 Matches: 339
 Percent Similarity: 99.12% Conservative: 0
 Best Local Similarity: 99.12% Mismatches: 3
 Query Match: 60.61% Indels: 1
 DB: 3 Gaps: 0

US-10-023-523-44 (1-546) x AAF21339 (1-1793)

QY 205 AsnSerGlnLysGlnMetLysLeuLeuGlnLysGlnSerGlnLeuValGlnLys 224
 DB 1791 AATTCACAGACCATGAAGCTCTCAGAAAAGCAGAGCCAGCTGGTGCAGAGAG 1732
 QY 225 AspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeu 244
 DB 1731 GACCACCTGGCGGTGAGCAGACAGAGCC-GTCTTGGCCCGCAGCAAGCTTGAGAGCCTA 1673
 QY 245 CysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGlyValGlnArgAlaArg 264
 DB 1672 TGCCGTGAGCTGCACGGGCACAAACCGCTCCCTCAAGGAAGAAGGTGTGCACGGCGCCGG 1613
 QY 265 GluGluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIle 284
 DB 1612 GAGGAGGAGGAGAGCCAGAGGTGACCTTCGACCTTCAGGTGACACTGAATGACATT 1553
 QY 285 GlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGlnAsnMetGlu 304
 DB 1552 CAGCTGCAGATGGAACACAGACACATGAGCGCAACTCCAAGCTGCGCAAGAGACATGGAG 1493
 QY 305 LeuAlaGluArgLeuLysLysLeuIleGlnGlnTyrGluLeuArgGluGluHisIleAsp 324
 DB 1492 CTGGCTGAGAGGCTCAAGAGCTGATTGAGCAGTATGAGTGGCGGAGGAGCATATCGAC 1433
 QY 325 LysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAla 344
 DB 1432 AAAGCTTCAACACACAGAGACCTACACAGCAGCTGGTGGATGCCAAGCTCCAGCAGGCC 1373
 QY 345 GlnGluMetLeuLysGluAlaGluArgHisGlnArgGluLysAspPheLeuLys 364
 DB 1372 CAGGAGATGCTAAAGAGGCGAGAGCGGCACAGCGCGGAGAGAGATTTTCTCTCTGAAA 1313
 QY 365 GluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlnThrHisLeuLys 384
 DB 1312 GAGGAGTAGAGTCCACAGAGGATGTGTGAGCTGATGAGCAGCAGAGACCCACTGAG 1253
 QY 385 GlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSer 404
 DB 1252 CAACAGCTTGCCCTATACACAGAGAAGTTTGAGGAGTTTCCAGAAACACACTTCCAAAGC 1193
 QY 405 SerGluValPheThrThrPheLysGlnGlnMetGluLysMetThrLysLysIleLysLys 424
 DB 1192 ACGGAGGTATTCACCACTTCAAGCAGGAGATGGAAAGATGACTAGAGATCAAGAG 1133

QY 425 LeuGluLysGluThrMetTyrArgSerArgTTPGluSerSerAsnLysAlaLeuLeu 444
 DB 1132 CTGGAGAAGAAACACCATGTACCGTCCCGTGGGAGAGCAGCAACAAGGCCCTGCTT 1073
 QY 445 GluMetAlaGluGluLysThrValArgAspLysGluLeuGluGlyLeuGluValLysIle 464
 DB 1072 GAGATGGCTGAGGAGAAACAGTCCGGGATTAAGAACTGGAGGGCTTCAGGTAAATAATC 1013
 QY 445 GlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArg 484
 DB 1012 CAACGGCTGGAGAACTGTGGGGCACTGCAGACAGAGCGCAATGACCTGAACAAGAGG 953
 QY 485 ValGlnAspLeuSerAlaGlyGlnGlySerLeuThrAspSerGlyProGluArgArg 504
 DB 932 GTACAGGACCTGAGTCTGGTGGCCAGGGCTCCCTCACTGACAGTGGCCCTGAGAGGAGG 893
 QY 505 ProGluGlyProGlyValAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyr 524
 DB 892 CCAGAGGGGGCTGGGGCTCAAGCACCCAGCTCCCGGAGGTCAACAGAGCGCTTGCTTAC 833
 QY 525 ProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAla 544
 DB 832 CCAGAGGACCGGACAGACAGATCAGGCCAGACTGGGCTTCAGAGCCCACTCCGCC 773
 QY 545 ArgAla 546
 DB 772 AGGGCC 767

RESULT 17

ABZ97034/c

ID ABZ97034 standard; DNA; 1793 BP.

XX AC ABZ97034;

DT 17-OCT-2003 (first entry)

XX DE Human nucleic acid sequence.

XX KW Human; antisense; lung dysfunction; nasal airway dysfunction;
 KW antinflammatory steroid; ubiquinone; antinflammatory; antiallergic;
 KW antinflammatory; hypotensive; immunosuppressive; cytostatic; gene therapy;
 KW antisense gene therapy; respiratory; lung; adenosine sensitivity;
 KW adenosine receptor; bronchodilation; bronchoconstriction; lung allergy;
 KW lung inflammation; respiratory disease; ds.

XX OS Homo sapiens.

XX WO200285308-A2.

XX PD 31-OCT-2002.

XX PF 23-APR-2002; 2002WO-US013135.

XX PR 24-APR-2001; 2001US-0286137P.

XX PA (EPG-) EPGENESIS PHARM INC.

XX PI Nyce JW, Li Y, Sandrasagra A, Katz E, Pabalan J, Aguilar D;

XX PI Miller S, Tang L, Shahabuddin S;

XX DR WPI; 2003-229219/22.

PT Pharmacological composition for treating ailments associated with impaired
 PT respiration, has oligo(s) antisense to specific gene(s) or its
 PT corresponding RNAs, and glucocorticoid or non-glucocorticoid steroid or
 PT ubiquinone.

XX PS Disclosure; SEQ ID NO 12276; 872pp; English.

XX CC The invention relates to a novel pharmaceutical composition, which has a
 CC first active agent comprising an oligonucleotide antisense to the
 CC initiation codon, coding region, 5' or 3' and genomic flanking regions,
 CC 5' and 3' intron-exon junctions, or regions within 2-10 nucleotides of

CC junctions of genes encoding a polypeptide associated with lung and/or
 CC nasal airway dysfunction and a second active agent comprising an
 CC antinflammatory steroid and ubiquinone. A composition of the invention
 CC has antinflammatory, antiallergic, antinflammatory, hypotensive,
 CC immunosuppressive, and cytostatic activity. The composition may have a
 CC use in antisense gene therapy. The composition is useful for treating or
 CC preventing a respiratory, lung or malignant disease or condition, also
 CC for enhancing the prophylactic or therapeutic respiratory effect of an
 CC antinflammatory steroid in a subject, for reducing or depleting levels
 CC of, or reducing sensitivity to adenosine, reducing levels of adenosine
 CC receptor, producing bronchodilation, increasing levels of ubiquinone or
 CC lung surfactant in a subject's tissue, or treating bronchoconstriction,
 CC lung inflammation, lung allergies, or a respiratory disease or condition.
 CC Note: The sequence data for this patent is not represented in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 1793 BP; 339 A; 546 C; 444 G; 464 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,01e-89 Length: 1793
 Score: 1688.00 Matches: 339
 Percent Similarity: 99.12% Conservative: 0
 Best Local Similarity: 99.12% Mismatches: 3
 Query Match: 60.61% Indels: 1
 DB: 7 Gaps: 0

US-10-023-523-44 (1-546) x ABZ97034 (1-1793)

QY 205 AsnSerGlnLysGlnMetLysLeuLeuGlnLysGlnSerGlnLeuValGlnGluLys 224
 DB 1791 AATTACAGAGCAGATGAAGCTCCACAGAAAAGCAGACCCAGCTGGTGCAGAGAAG 1732
 QY 225 AspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeu 244
 DB 1731 GACCACCTGCGGTGAGCACAGCAGCC-GTCTGCGCCGACCAAGCTTGAGAGCCTA 1673
 QY 245 CysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGluValGlnArgAlaArg 264
 DB 1672 TGCGGTGAGTGACGCGGCACCAACCGCTCCCTCAAGAGAAAGGTGTGCAGCGGCCCG 1613
 QY 265 GluGluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIle 284
 DB 1612 GAGAGAGAGAGAGAGCGCAGAGAGTGACCTCCGACCTTCGAGTGACATGATGACATT 1553
 QY 285 GlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGlu 304
 DB 1552 CAGCTGCAGATGAACACAGCACAATGAGCGCACTCCAAGCTGCGCCCAAGAGAACATGGAG 1493
 QY 305 LeuAlaGluArgLeuLysLysLeuLysGluGluGluGluGluGluHisIleAsp 324
 DB 1492 CTGGCTGAGAGGCTCAAGAAAGCTGATTGACAGTATGAGCTGCGGAGGAGCATATCGAC 1433
 QY 325 LysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnAla 344
 DB 1432 AAGTCTTCAACACACAGGACCTACACAGCAGCTGGTGCATGCCAAGCTCCAGCAGGCC 1373
 QY 345 GlnGluMetLysGluAlaGluArgHisGlnArgGluLysAspPheLeuLys 364
 DB 1372 CAGGAGATGCTAAGAGGAGGAGAGCGCCACCGAGGAGAGGATTTCCTCTCTCTCTCT 1313
 QY 365 GluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGluThrHisLeuLys 384
 DB 1312 GAGCAGTATAGTCCACAGAGATGTGTGAGCTGATGNAGCAGCAGAGAGACCCACCTGAAG 1253
 QY 385 GlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSer 404
 DB 1252 CAACAGCTTGCCTATACACAGAGAGTTTTCAGGAGTTTTCAGAGAACACACTTTCCAAAGC 1193
 QY 405 SerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLys 424
 DB 1192 AGCAGGTATTTCACCATTCAGCAGAGAGATGGAAGAAAGATGACTAAGAGATCAGAGAG 1133

QY 425 LeuGluLysGluThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeu 444
 DB 1132 CTGGAGAAAGAACACCACTATGACCGGTCCCGTGGAGAGAGCAACAAGCCCTGCTT 1073
 QY 445 GluMetAlaGluGluLysThrValArgAspLysGluLeuGluGluLysGlnValLysIle 464
 DB 1072 GAGATGGCTGAGGAGAAACAGTCCGGATTAAGAACTGGAGGGCTGCGAGTAAATC 1013
 QY 465 GlnArgLeuGluLysLeuCysArgAlaLeuGlnInThrGluArgAsnAspLeuAsnLysArg 484
 DB 1012 CAACGGCTGAGAAAGCTGCGCGGCACCTGCAGACAGAGCGCAATGACCTGAACAAGAGG 953
 QY 485 ValGlnAspLeuSerAlaGlyGlnGlySerLeuThrAspSerGlyProGluArgArg 504
 DB 952 GTACAGAGCTGAGTCTGGTGGCCAGGGCTCCCTACTGACAGTGGCCCTGAGAGGAGG 893
 QY 505 ProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyr 524
 DB 892 CCAGAGGGGCTGGGCTCAAGCACCCAGCTCCCGGAGGTCACAGAGGGCTTGCTAC 833
 QY 525 ProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAla 544
 DB 832 CCAGGAGCACCGAGCACAGAGCATCAGGCCAGACTGGGCTTCAAGAGGCCACCTCCGCC 773
 QY 545 ArgAla 546
 DB 772 AGGGCC 767

RESULT 18
 ABZ97033/C
 ID ABZ97033 standard; DNA; 1793 BP.
 AC ABZ97033;
 XX
 DT 17-OCT-2003 (first entry)
 XX
 DE Human nucleic acid sequence.
 XX
 KW Human; antisense; lung dysfunction; nasal airway dysfunction;
 KW antiinflammatory steroid; ubiqunone; antiinflammatory; antiasthmatic;
 KW antiasthmatic; hypotensive; immunosuppressive; cytostatic; gene therapy;
 KW antisense gene therapy; respiratory; lung; adenosine sensitivity;
 KW adenosine receptor; bronchodilation; bronchoconstriction; lung allergy;
 KW lung inflammation; respiratory disease; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200285308-A2.
 XX
 PD 31-OCT-2002.
 XX
 PF 23-APR-2002; 2002WO-US013135.
 XX
 PR 24-APR-2001; 2001US-0286137P.
 XX
 PA (EPITG-) EPIGENESIS PHARM INC.
 XX
 PI Nyce JW, Li Y, Sandrasagra A, Katz E, Pabalan J, Aguilar D;
 PI Miller S, Tang L, Shahabuddin S;
 XX
 DR WPI; 2003-229219/22.
 XX
 PT Pharmaceutical composition for treating ailments associated with impaired
 PT respiration, has oligo(s) antisense to specific gene(s) or its
 PT corresponding RNAs, and glucocorticoid or non-glucocorticoid steroid or
 PT ubiqunone.
 XX
 PS Disclosure; SEQ ID NO 12275; 872pp; English.
 XX
 CC The invention relates to a novel pharmaceutical composition, which has a
 CC first active agent comprising an oligonucleotide antisense to the
 CC initiation codon, coding region, 5' or 3' and genomic flanking regions,
 CC 5' and 3' intron-exon junctions, or regions within 2-10 nucleotides of

CC junctions of genes encoding a polypeptide associated with lung and/or
 CC nasal airway dysfunction and a second active agent comprising an
 CC antiinflammatory steroid and ubiqunone. A composition of the invention
 CC has antiinflammatory, antiasthmatic, antiasthmatic, hypotensive,
 CC immunosuppressive, and cytostatic activity. The composition may have a
 CC use in antisense gene therapy. The composition is useful for treating or
 CC preventing a respiratory, lung or malignant disease or condition, also
 CC for enhancing the prophylactic or therapeutic respiratory effect of an
 CC antiinflammatory steroid in a subject, for reducing or depleting levels
 CC of, or reducing sensitivity to adenosine, reducing levels of adenosine
 CC receptor, producing bronchodilation, increasing levels of ubiqunone or
 CC lung surfactant in a subject's tissue, or treating bronchoconstriction,
 CC lung inflammation, lung allergies, or a respiratory disease or condition.
 CC Note: The sequence data for this patent is not represented in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences

XX
 SQ Sequence 1793 BP; 339 A; 546 C; 444 G; 464 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,01e-89 Length: 1793
 Score: 1688.00 Matches: 339
 Percent Similarity: 99.12% Conservative: 0
 Best Local Similarity: 99.12% Mismatches: 3
 Query Match: 60.61% Indels: 1
 DB: 7 Gaps: 0

US-10-023-523-44 (1-546) x ABZ97033 (1-1793)

QY 205 AsnSerGlnLysGlnMetLysLeuGlnLysLysGlnSerGlnLeuValGlnGluLys 224
 DB 1791 AATTCAAGAGCAGATGAAGTCTCTACAGAAAAGCAGAGCAGCTGTGCAAGAGAAG 1732
 QY 225 AspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeu 244
 DB 1731 GACCACCTGGCGGTGAGCAGACCAAGCC-GTCTGGCCCGCAGCAGAGTTTGAAGACCTA 1673
 QY 245 CysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGluValGlnArgAlaArg 264
 DB 1672 TGGCGTGAGCTGCACGGGCACACCCCTCCCTCAAGGAAAGAGGTGTGCGCGGGCCGG 1613
 QY 265 GluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIle 284
 DB 1612 GAGCAGAGAGAGAGCAGCGCAAGGAGGTGACTCGCACTTCAGGTGACACTGAATGACATT 1553
 QY 285 GlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGlnAsnMetGlu 304
 DB 1552 CAGCTGCAGATGGAACAGACACATGAGCGCAACTCCCAAGCTGCGCCACAGAACTGGAG 1493
 QY 305 LeuAlaGluArgLeuLysLysLysLeuIleGluGlnTyrGluLeuArgGluLysIleAsp 324
 DB 1492 CTGGCTGAGAGGCTCAAGAGGCTGATTGAGCAGTATGAGCTGCGCGAGGAGCATATCGAC 1433
 QY 325 LysValPheLysHisLysAspLeuGlnGlnLeuValAspAlaLysLeuGlnAla 344
 DB 1432 AAAGTCTTCAACACACAGAGACCTACACAGCAGCTGTGTGATGCCAAGCTCCAGCAGGCC 1373
 QY 345 GlnGluMetLysGlnAlaGluGluArgHisGlnArgGluLysAspPheLeuLys 364
 DB 1372 CAGGAGATGCTAAAGAGGAGCAGAGAGCGCACCCAGCGGAGAGGATTTCTCTCGTAAA 1313
 QY 365 GluAlaValCysSerGlnArgMetCysGluLeuMetLysGlnGlnGlnThrHisLeuLys 384
 DB 1312 GAGCAGTAGAGTCCAGAGAGTGTGTGAGCTGTATGAGCAGCAGCAGAGACCCACCTGAG 1253
 QY 385 GlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSer 404
 DB 1252 CAACAGCTTGCCCTATACACAGAGAGAGTTTGAAGAGTTCACAGACACACTTCCAAAGC 1193
 QY 405 SerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLys 424
 DB 1192 AGCAGGTATTCCACCACTTCAAGCAGAGGAGATGGAAGATGACTAAGAATCAAGAAG 1133

QY 425 LeuGluLysGluThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeu 444
 Db 1132 CTGGAGAAAGAACCAACCATGTACCGTCCCGGTGGGAGAGCAGCAACCAAGGCCCTGTT 1073
 QY 445 GluMetAlaGluLysThrValArgAspLysGluLeuGluGlyLeuGlnValLysIle 464
 Db 1072 GAGATGGCTGAGGAGAAACAGTCCCGGATAAGAACTGGAGGCCCTGAGGTAAATC 1013
 QY 465 GlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArg 484
 Db 1012 CAACGGCTGGAGAGCTGTCCCGGCACCTGCAGACAGAGCGCAATGACCTGAACAGAGG 953
 QY 485 valGlnAspLeuSerLagGlyGlnGlySerLeuThrAspSerGlyProGluArgArg 504
 Db 952 GTACAGGACCTGAGTGTGGTGGCCAGGGCTCCCTCACGTGACAGTGGCCCTGAGAGGAG 893
 QY 505 ProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyr 524
 Db 892 CCAGAGGGGCTGGGGCTCAAGCACCCAGCTCCCGCAGGTCACAGAGGGCTTGCTAC 833
 QY 525 ProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAla 544
 Db 832 CCAGGACCCGAGCAGACAGAAACATCAGGCCAGACTGGGCTCAGAGGCCACCTCCGCC 773
 QY 545 ArgAla 546
 Db 772 AGGGCC 767

RESULT 19

AAQ66987/C
 ID AAQ66987 standard; cDNA to mRNA; 1854 BP.

XX AC AAQ66987;

XX 25-MAR-2003 (revised)

DT 19-MAR-1995 (first entry)

DE Interleukin 14 gene.

XX IL-14; B-cell growth factor; proliferation; differentiation; replication;
 KW non-Hodgkin's lymphoma; systemic lupus erythematosus;
 KW systemic necrotizing vasculitis; transplant; ss.

OS Homo sapiens.

XX Key Location/Qualifiers
 FH 73..1770
 FT /*tag= a

XX WO9416074-A2.

XX 21-JUL-1994.

XX 18-JAN-1994; 94WO-US001101.

XX 15-JAN-1993; 93US-00005156.

XX (UYWA-) UNIV WASHINGTON ST LOUIS.
 XX (USH) US DEPT HEALTH & HUMAN SERVICES.
 XX (TEXA) UNIV TEXAS SYSTEM.

XX Ambrus JL, Fauci AS, Ford RJ;

XX WPI, 1994-249221/30.

XX P-PSDB; AAR55800.

XX High molecular weight B cell growth factor - able to stimulate B cell
 PT proliferation and inhibit B cell differentiation, useful to treat e.g.
 PT systemic lupus erythematosus.

XX Claim 4; Page 75; 95pp; English.

XX The sequence is that of a gene encoding a high molecular weight B cell

CC growth factor which stimulates B cell proliferation and inhibits B cell
 CC differentiation, termed interleukin 14. IL-14 can be used to interfere
 CC with the replication of B cells and to treat B cell non-Hodgkin's
 CC lymphoma, systemic lupus erythematosus and systemic necrotizing
 CC vaculitis. (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 1854 BP; 350 A; 559 C; 456 G; 489 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4.09e-84 Length: 1854
 Score: 1597.00 Matches: 334
 Percent Similarity: 97.66% Conservative: 0
 Best Local Similarity: 97.66% Mismatches: 5
 Query Match: 57.34% Indels: 5
 DB: 2 Gaps: 0

US-10-023-523-44 (1-546) x AAQ66987 (1-1854)

QY 208 LysGlnMetLysLeuLeuGlnLysGlnSerGlnLeuValGlnGluLysAspHisLeu 227
 Db 1798 AAGCAGATGAAGTCTCTACAGAAAAAGCAGAGCCAGCTGGTGCAGAGAGACACCTG 1739
 QY 228 ArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGlu 247
 Db 1738 CGCGGTGAGCACAGCAAGCC-GTCTGCCCGCAGCAAGCTTGAGAGCCTATGCCGTGAG 1680
 QY 248 LeuGlnArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGluGluGlu 267
 Db 1679 CTGCACGGGCACACCGCTCTCTCAGGAAGAGGTGTCAGCGGCCCGGAGGAGGAG 1620
 QY 268 GluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGln 287
 Db 1619 GAGAAAGCGCAAGAGGTGACCTCGCACCTCCAGGTGACACCTGAATGACATTCAGTTCGAG 1560
 QY 288 MetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGluLeuAlaGlu 307
 Db 1559 ATGGAAACAGCAACATGAGCGCAACTCCAACTGCCCCAGAGAAACATGGAGGTGGCTGAG 1500
 QY 308 ArgLeuLysLysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAspLysValPhe 327
 Db 1499 AGGCTCAAGAACTGATTGAGCAGTATGAGTGCAGGAGGACATATCGACAAGTCTTC 1440
 QY 328 LysHisLysAspLeuGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMet 347
 Db 1439 AAACCAAGAGACCTACACAGCAGCTGGTGGATCCAAAGCTCCAGCAGGCCCAGGAGATG 1380
 QY 348 LeuLysGluAlaGluArgHisGlnArgGluLysAspPheLeuLysGluAlaVal 367
 Db 1379 CTAAGGAGCGCAGAGAGCGGCCACCGCGGAGAGAGGATTTTCTCTGAAAGAGGCAGTT 1320
 QY 368 GluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLeuLysGlnGlnLeu 387
 Db 1319 GAGTCCAGAGGATGTGTGAGCTGATGAAGCAGCAGAGACCCACCTGAAGCAACAGCTT 1260
 QY 388 AlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluVal 407
 Db 1259 GCCCTATACAGAGAAGTTTGAGGAGTTCCAGACACACTTTCGAAAAGCAGCGAGGTA 1200
 QY 408 PheThrThrPheLysGlnGluMetGluLysMetThrLysLysIleLysLysLeuGluLys 427
 Db 1199 TTCACCACTTCAAGCAGGAGATGGAAGATGACTAAGAAGATCAAGAAGCTGGAGAA 1140
 QY 428 GluThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeuMetAla 447
 Db 1139 GAAACCACTATGACCGTCCCGGTGGGAGAGCAGCAACAGGCCCTGCTTGGATGGCT 1080
 QY 448 GluGluLysThrValArgAspLysGluLeuGluGlyLeuGlnValLysIleGlnArgLeu 467
 Db 1079 GAGAGAAACAGTCCGGGTATAGAACTGGAGGGCTTCAGGTAAAAATCCAAAGCTG 1020
 QY 468 GluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgValGlnAsp 487
 Db 1019 GAGAACTGTGCCGGCCTGACAGAGAGCGCAATGACCTGAACAGAGGGGTACAGGAC 960

QY 488 LeuSerAlaGlyGlnGlySerLeuThrAspSerGlyProGluArgArgProGlu-G1 507
 Db 959 CTGAGTGTCTGGCCAGGGCTCTCTACTGACAGTGGCCCTGAGAGAGGCCAGAGGGG 900
 QY 507 yProGlyAla-GlnAlaProSerSerProArgValThrGluAlaProCysTyrProGlyA 527
 Db 899 CACTGGGCGATCAAGCACCAGCTCCCGGAGGTCACAGAGCGCTTGCTACCCAGAG 840
 QY 527 la-ProSerThrGluAlaSerGlnThrGlyProGlnGluProThrSerAlaArgAla 546
 Db 839 CAACCGAGCAGAGAGCATCAGGCGAGACTGG-CCTCAGAGAGCCACCTCGCCAGGGCC 781

RESULT 20

AAH15042

ID AAH15042 standard; cDNA; 2045 BP.

AC AAH15042;

XX 26-JUN-2001 (first entry)

XX Human cDNA sequence SEQ ID NO:13025.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-00116126.

XX 29-JUL-1999; 99JP-00248036.

XX 27-AUG-1999; 99JP-00300253.

XX 11-JAN-2000; 2000JP-00118776.

XX 02-MAY-2000; 2000JP-00163767.

XX 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

PS Claim 8; SEQ ID NO 13025; 2537pp + Sequence listing; English.

XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92445 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent

CC oligonucleotides, all of which are used in the exemplification of the CC present invention

XX SQ Sequence 2045 BP; 672 A; 421 C; 506 G; 446 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5-59e-66 Length: 2045
 Score: 1286.00 Matches: 284
 Percent Similarity: 67.4% Conservatives: 71
 Best Local Similarity: 53.9% Mismatches: 132
 Query Match: 46.1% Indels: 39
 DB: 4 Gaps: 8

US-10-023-523-44 (1-546) x AAH15042 (1-2045)

QY 28 GluGlyAlaGlnGluArgProSerGlnAlaAlaProAlaValGluAlaGluGlyProGly 47
 Db 70 GAGGCGAGCGCGGGAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 129
 QY 48 SerSerGlnAlaProArgLys-----ProGluGlyAlaGlnAlaArgThrAlaGlnSer 65
 Db 130 CGCGGAGCGAGCG 189
 QY 66 GlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGluAspIleLeuSerThr 85
 Db 190 GGGCTAGGGGTGAAGAGCAGATATGTTGTAACTCTCAATCAATGATATCTTCAACAT 249
 QY 86 TyrCysValAspAsnGlnGlyClyProGlyGluAspGlyAlaGlnGlyGluProAla 105
 Db 250 CAA-----GGCTCAATTTGGTGGCACAAGTAAAGCAAGCATTCATTGGAAGAGATGAA 303
 QY 106 GluProGluAspAlaGluSerArgThrTyrValAlaArgAsnGlyGluProGluPro 125
 Db 304 GGCAGTGACTTTATAACAGACAGACAGGAATTTGGTGGC-----CCA 345
 QY 126 ThrProValValAsnGlyGlyGluProSerLysGlyAspProAsnThrGluGluIle 145
 Db 346 GCATACTGCACGCAAGAATCAAGAGAGGAAATCCCTGGGGA-----GAAGCT 393
 QY 146 ArgGlnSerAspGluValGlyAspArgAspHisArgArgProGlnGluLysLysAla 165
 Db 394 CGAACAGATCCCCCTGATGGTGCAGCAGATTCAGAGTCCACAGGAAACAAAGAAACT 453
 QY 166 LysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrPro 185
 Db 454 -----TTAGAAAGAGAGTTTATTACTGATCAAGCCCTTAAACACCCCTTCAACCCA 507
 QY 186 GluGluLysLeuAlaAlaLeuCysLysLysTyrAlaGluLeuLeuGluGluHisArgAsn 205
 Db 508 GAGGAGAGCTGGCAGCTCTCTGTAAAGAAATATCTGATCTTCTGGAGGAGCAGAGT 567
 QY 206 SerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAsp 225
 Db 568 GTTCAGAGCAATGAAGATCCTGCAGAGAAGCAAGCCAGCATTTGTGAAGAGAAAGTT 627
 QY 226 HisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCys 245
 Db 628 CACTTGCAGATGACATAGCAGGCTATCTTGGCAAGAGCAAGCTAGATCTTCTTTC 687
 QY 246 ArgGluLeuGlnArgHisAsnArgSerLeuLysGluGlyValGlnArgAlaArgGlu 265
 Db 688 AGAGAACTTCAGCTCAATAAGACGCTTAAAGAGGAAATAATGCAGAGGAGCAGAGAG 747
 QY 266 GluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGln 285
 Db 748 GAAGAGAACACGCTAAGAGAGCAACTGCACATTTCCAGATTACCTTAGATGAATCAA 807
 QY 286 LeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGluLeu 305
 Db 808 GCCCAGCTGGAGCAGCATGACATCCACACGCCAACTCCGACAGGAAACATTTAGCTG 867
 QY 306 AlaGluArgLeuLysLysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAspLys 325

Db 868 GGGGAGAGCTAAAGAGCTCATCGAACAGTACGACGTGCGGAGAGACACATTGATAAG 927
Qy 326 ValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGln 345
Db 928 GTGTTCAACCTAAGGAACTGCACAGCAGCTCGTGGATGCCAACTGCAGCAACGACA 987
Qy 346 GluMetLeuLysGluAlaGluGluArgHisGlnArgGluLysAspPheLeuLeuLysGlu 365
Db 988 CAACCTGATAAAGAGCTGATGAAAGAACATCAGAGAGAGAGATTTTATTAAAGAA 1047
Qy 366 AlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlnThrHisLeuLysGln 385
Db 1048 GCGACAGATGAGGCAACAAATACGAAATGAAAGCAGAGAGATACACTTAAACAG 1107
Qy 386 GlnLeuAlaLeuThrThrGluLysPheGluGluPheGlnAenThrLeuSerLysSerSer 405
Db 1108 CAGCTTCTCTTATATGATAAGTTTGAAGAAATCCAGACTACCTGCGCAAGAAAGCAAT 1167
Qy 406 GluValPheThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLysLys 425
Db 1168 GAACCTGTTTACACCTTCAGACAGGAAATGGAAGATGACAAAGAAATTAATAAACTG 1227
Qy 426 GluLysGluThrThrMetTyArgSerArgTyrGluSerSerAsnLysAlaLeuLeuGlu 445
Db 1228 GAAAAGAAACAAATATTTGGGTACCAATGGGAACAAATAATAAGCACTTCTGCAA 1287
Qy 446 MetAlaGluGluLysThrValArgAspLysGluLeuGluGluLysLysLysLysLys 465
Db 1288 ATGGCTGAAGAGAAACAGTCCGTGATAAGAGTACAAAGGCCCTTCAAAATAAACTGAA 1347
Qy 466 ArgLeuGluLysLysCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgVal 485
Db 1348 CGTTAGAGAGCTGTGAGGGCTCTTCAACAGAAAGATGAGCTCAATGAGAAGGTG 1407
Qy 486 Gln----- 488
Db 1408 GAAGTCTCTGAAAGAGCAGGTATCCATCAAGCGGCCATCAAGCGCGCAACAGGGATT 1467
Qy 489 SerAlaGlyGlyGlnGlySerLeuThr-----AspSerGlyProGluArgArgProGlu 506
Db 1468 GCAACACCTGTGATGCACCCCTGTACTCCCTGGATTCTCAAGAGGCTGAACACTTCC 1527
Qy 507 GlyPro-----GlyAlaGlnAlaProSer-SerProArgValThrGluAlaProCys 523
Db 1528 TCGAAAGAGCCCTGGAGCGCACCTGGAGGCTGAGCCCAAGAGTTCAGAGAAGCGCTGTG 1587
Qy 523 sTyrProGlyAlaPro 528
Db 1588 CAAAAGCCCCCGTCCA 1603
RESULT 21
ID ABK70301
XX AC ABK70301;
XX DT 15-JUL-2002 (first entry)
XX DE Human lung cancer associated full length cDNA DMSM-181.
XX KW Human; ss; gene; lung cancer; cytostatic; tumour; vaccine.
XX OS Homo sapiens.
XX PN WO200224057-A2.
XX PD 28-MAR-2002.
XX PF 20-SEP-2001; 2001WO-US042232.
XX PR 22-SEP-2000; 2000US-0234837P.
XX PR 10-OCT-2000; 2000US-023940P.
XX PR 29-JUN-2001; 2001US-0301928P.

XX PA (CORI-) CORIXA CORP.
XX PI Benson DR, Mohamath R, Lodes MJ;
XX DR WPI; 2002-372001/40.
XX PT New tumor lung proteins and nucleic acids encoding the proteins, useful
PT as vaccines and for treating, preventing, diagnosing or monitoring lung
PT cancer.
XX PS Claim 1; Page 175-176; 189pp; English.
XX CC The invention relates to an isolated polynucleotide comprising a sequence
CC selected from 183 human DNA sequences (appearing as ABK70130-ABK70312),
CC or their fragments, homologues, variants or complements and their encoded
CC polypeptides. Also included are an expression vector comprising the
CC polynucleotide operably linked to an expression control sequence; a host
CC cell transformed or transfected with an expression vector of; an isolated
CC antibody or its antigen-binding fragment that specifically binds to the
CC polypeptide; a method for detecting the presence of a cancer in a patient
CC ; a fusion protein comprising at least the polypeptide; an
CC oligonucleotide that hybridises to the polynucleotide under moderately
CC stringent conditions; a method for stimulating and/or expanding T cells
CC specific for a tumor protein; an isolated T cell population comprising a first
CC cells prepared from the method of above; a composition comprising a first
CC component consisting of carriers and immunostimulants, and a second
CC component selected from the polynucleotides, proteins, antibodies, fusion
CC proteins, T cell populations and antigen presenting cells expressing the
CC polypeptide; methods for stimulating an immune response or treating
CC cancer in a patient by administering the composition and diagnostic kits
CC comprising at least one of the oligonucleotide of, or an antibody and a
CC detection reagent consisting of a reporter group. The polypeptides and a
CC polynucleotides are useful as vaccines for the treatment or prevention of
CC lung cancer, and for diagnosis and monitoring of such cancer. The
CC polynucleotide and antigen presenting cells can be used to
CC stimulate or expand T cells specific for a tumorous protein. The
CC polynucleotides may be used as probes or primers for nucleic acid
CC hybridisation, and in the preparation of ribozyme molecules for
CC inhibiting expression of tumor polypeptides and proteins in tumour
CC cells. The present sequence is one of the 183 lung cancer associated
CC polynucleotides
XX SQ Sequence 2045 BP; 672 A; 421 C; 506 G; 446 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5,59e-66 Length: 2045
Score: 1286.00 Matches: 284
Percent Similarity: 67.49% Conservative: 71
Best Local Similarity: 53.99% Mismatches: 132
Query Match: 46.18% Indels: 39
DB: 6 Gaps: 6

US-10-023-523-44 (1-546) X ABK70301 (1-2045)
Qy 28 GluGlyAlaGlnGluArgProSerGlnAlaAlaProAlaValGluAlaGluGlyProGly 47
Db 70 GAGGCAAGCGCGGGAAGAGCGCGCGCGCGGAGAGCGCGCGCGCGCGCGCGCGCGGA 129
Qy 48 SerSerGlnAlaProArgLys-----ProGluGlyAlaGlnAlaArgThrAlaGlnSer 65
Db 130 CGCGAGCG 189
Qy 66 GlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGluAspLeuSerThr 85
Db 190 GGGCTAGGGGTGAAGCAGATATGTTGTGTAACCTCAATCAATCATATCTTCAACAT 249
Qy 86 TyrCysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAlaGlnProAla 105
Db 250 CAA-----GGCTCAAAATTTGTGTGCGACCAAGTAACAAGCATTCATTGAAGAGTGA 303
Qy 106 GluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsnGlyGluProGluPro 125

Db 304 GCGAGTACTTTATACAGAACAGGAAATTTGGTGAGC-----CCA 345
 QY 126 ThrProValValAsnGlyGluLysGluProSerLysGlyAspProAsnThrGluGluLeu 145
 Db 346 GCATACTGCGACGCAAGAAATCAAGAGGAAATCCCTGGGGGA-----GAAGCT 393
 QY 146 ArgGlnSerAspGluValGlyAspArgAspHisArgArgProGlnGluLysLysLysala 165
 Db 394 CGAACAGATCCCTGATGTCAGCAGATTCAGAGTGCACAGAGCAAAAGAAAAACT 453
 QY 166 LysGlyLeuGlyLysGluLeuThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrPro 185
 Db 454 -----TTAGGAAAAAGAGTTTATTACTGATGCAAGCCCTAAACACCTTTCAACCCCA 507
 QY 186 GluGluLysLeuAlaAlaLeuCysLysLysLysLysLysLysLysLysLysLysLysLys 205
 Db 508 GAGGAGAGAGCTGGAGCTCTCTGTAGAAATATGCTGATCTCTGGAGAGAGCAGGAGT 567
 QY 206 SerGlnLysGlnMetLysLeuLeuGlnLysLysLysLysLysLysLysLysLysLys 225
 Db 568 GTTCAGAGCAATGAGATCTGAGAGAGCAGCCAGCATTTGTGAAGAGAAAGTT 627
 QY 226 HisLeuArgGlyGluHisSerLysLysLysLysLysLysLysLysLysLysLysLysLys 245
 Db 628 CACTTGCAGAGTCAACATAGCAAGCTATCTTGGCAAGAGCAAGCTAGAAATCTTTTGC 687
 QY 246 ArgGluLeuGlnArgHisAsnArgSerLysLysGluGluGlyValGlnArgAlaArgGlu 265
 Db 688 AGAGAACTTCACGGTCACANTAGACGTTAAGAGAGAAATATGCACAGCAGCAGAG 747
 QY 266 GluGluGluLysArgGluValThrSerHisPheGlnValThrLeuAsnAspIleGln 285
 Db 748 GAAGAAGAACGACGTAAGCAACTGCAACCACTCCAGATTTCCAGATTACCTTAGATGAATTC 807
 QY 286 LeuGlnMetGluGlnHisAsnGluArgAsnSerLysLysLysLysLysLysLysLysLys 305
 Db 808 GCCAGCTGGAGCAGATGACATCCACACCGCAACTCCGACAGAGAAACATTTGAGCTG 867
 QY 306 AlaGluArgLeuLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 325
 Db 868 GGGGAGAGACTAAGAGCTCATCGACAGTACGCACTGAGGAGAGACACATTTGATAAG 927
 QY 326 ValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGln 345
 Db 928 GTGTTCAAAACGTAAGCAACTGCAACAGCAGCTCGTGGATGCCAACTCCAGCAACCGACA 987
 QY 346 GluMetLeuLysGluAlaGluArgHisGlnArgGluLysAspPheLeuLysLysGlu 365
 Db 988 CAACTGATAAAGAGCTGATGAAAAATCATCAGAGAGAGAGAGATTTTATTAAAGAA 1047
 QY 366 AlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlnThrHisLysLysGln 385
 Db 1048 GCGACAGAAATCGAGGACCAAAATPACCAAAATGAAACAGCAGGAGTACAACTTAAACAG 1107
 QY 386 GlnLeuAlaLeuThrThrClyLysPheGluPheGlnAsnThrLeuSerLysSerSer 405
 Db 1108 CAGCTTTCTTTATATGATGAAGTTTGAAGATTTCCAGTACCATCGCAAAAGCAAT 1167
 QY 406 GluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLysLys 425
 Db 1168 GAACCTGTTTACAACTTCACAGCAGCAAAATGGAAGAGATGACAAAGAAATTTAAAAACTG 1227
 QY 426 GluLysGluThrThrMetTyArgSerArgTrpGluSerSerAsnLysAlaLeuLeuGlu 445
 Db 1228 GAAAGAAAGAAACAAATATTTGGCGTACCATAATGGGAAACAAATAATAGCACTTCTGCA 1287
 QY 446 MetAlaGluGluLysThrValArgAspLysGluLeuGluGlyLeuGlnValLysIleGln 465
 Db 1288 ATGGCTGAAGAGAGAAACAGTCCGTGATAAGAGTACAGGCCCTTCAATAAAACTGGA 1347
 QY 466 ArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgVal 485
 Db 1348 CGGTTAGAGAGCTGTGAGGCTCTTCAACAGAGAAATGAGCTCAATGAGAGAGTG 1407

QY 486 Gln-----aspLeu 488
 Db 1408 GAAGTCTCTGAAGAGCAGGATTCATCAAAAGCGGCCATCAAAAGCGCGCAACAGGATTTA 1467
 QY 489 SerAlaGlyGlyGlnGlySerLeuThr-----AspSerGlyProGluArgArgProGlu 506
 Db 1468 GCAACACTGCTGTATGTCAGCCCTGTACTGCTGCTGATTTCTCAAAAGAGCTGAACACTTCC 1527
 QY 507 GlyPro-----GlyAlaGlnAlaProSer-SerProArgValThrGluAlaProCys 523
 Db 1528 TCGAAAGAGCCCTGGAGCGCACCTGGAGCTGAGCCCAAGAGTCAGAGAAAGCGCTGTG 1587
 QY 523 sTyProGlyAlaPro 528
 Db 1588 CAAAGCCCCCTCCA 1603

RESULT 22

ACD13396
 ID ACD13396 standard; cDNA; 1618 BP.
 XX ACD13396;
 AC ACD13396;
 XX 13-AUG-2003 (first entry)
 DT
 DE Human DNA encoding a p53 modifier, SEQ ID 67.

Human; ss; gene; p53 modifier; cytostatic; cancer; cytostatic;
 antiangiogenic; antiapoptotic; p53 pathway; breast cancer; colon cancer;
 kidney cancer; lung cancer; ovarian cancer; angiogenesis; cell cycle;
 apoptotic disorder; cell proliferation disorder.

Homo sapiens.

WO200299122-A1.

12-DEC-2002.

03-JUN-2002; 2002WO-US017382.

05-JUN-2001; 2001US-0296076P.

10-OCT-2001; 2001US-0328605P.

15-FEB-2002; 2002US-0357253P.

(EXEL-) EXELIXIS INC.

Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;

WPI; 2003-1556859/15.

P-PSDB; ABO07222.

Identifying modulators of the p53 pathway for use in treating apoptotic or cell proliferation disorders, comprises screening for agents that modulate activity of a human ortholog of genes that modify the p53 pathway in *Drosophila*.

Example 2; Page 319-320; 678pp; English.

The invention relates to identifying (M1) a candidate p53 pathway modulating agent, by contacting an assay system comprising a purified HM polypeptide (human orthologue of genes that modify the p53 pathway in *Drosophila*) or nucleic acid with a test agent under conditions, where but for the presence of the test agent, the system provides a reference activity, and detecting a test agent-biased activity of the assay system. Also included are modulating (M2) a p53 pathway of a cell (comprising contacting a cell defective in p53 function with a candidate modulator that specifically binds to a HM polypeptide comprising an HM amino acid sequence, where p53 function is restored), modulating (M3) a p53 pathway in a mammalian cell (comprising contacting the cell with an agent that specifically binds an HM polypeptide or nucleic acid) and diagnosing (M4) a disease in a patient (comprising: (a) obtaining a biological sample from the patient; (b) contacting the sample with a probe for HM expression; (c) comparing the results with a control; and (d) determining

PR 04-JAN-2002; 2002US-0345008P.
 PR 18-MAR-2002; 2002US-0365645P.
 PR 10-MAY-2002; 2002US-0379887P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Honchell CD, Warren BA, Borowsky ML, Griffin JA, Li JX, Lee SY;
 PI Yue H, Forsythe IJ, Marquis JP, Gietzen KJ, Baughn MR, Tran UK;
 PI Lehr-Mason PM, Tang YT, Ramkumar J, Emerling BM, Lee EA, Elliott VS;
 PI Hafalia AJA, Duggan BM, Chawla NK, Kable AE, Chang H, Khare R;
 PI Becha SD, Jin P, Lee S;
 XX WPI; 2003-363137/34.
 DR P-PSDB; ABR43284.
 XX
 PT New human neurotrophin associated proteins (NTRN), useful for
 PT diagnosing, treating and preventing diseases or conditions associated
 PT with the aberrant NTRN expression e.g. cancer, AIDS, diabetes, epilepsy,
 PT or infections.
 XX
 PS Claim 5; Page 230-231; 240pp; English.
 CC ACC68979 to ACC69003 encode the human neurotrophin-associated
 CC proteins given in ABR43271 to ABR43295, designated NTRN-1 to NTRN-25
 CC (I). (I) have cytostatic, antiarteriosclerotic, anticonvulsant,
 CC neurotrophic, neuroprotective, cerebroprotective, anti-HIV, anti-allergic,
 CC anti-inflammatory, thymimetic and antidiabetic activities, and can be
 CC used in gene therapy. The NTRN polypeptides and polynucleotides are
 CC useful in diagnosing, treating and preventing diseases or conditions
 CC associated with the decreased expression or overexpression of NTRN, such
 CC as cell proliferative (e.g. cancer, atherosclerosis), neurological (e.g.
 CC epilepsy, Huntington's disease, stroke), immune/inflammatory (e.g. AIDS,
 CC allergies), developmental (e.g. hypothyroidism, Cushing's syndrome) or
 CC endocrine (e.g. diabetes) disorders, or infections. They are also useful
 CC in assessing the effects of exogenous compounds on the expression of
 CC nucleic acid and amino acid sequences of NTRN. The NTRNs or their
 CC fragments are useful in screening compounds for effectiveness as agonist
 CC or antagonist of the polypeptides, or in altering the expression of the
 CC target polynucleotide and compounds that specifically bind to or modulate
 CC the activity of the polypeptide
 XX
 SQ Sequence 4174 BP; 1331 A; 924 C; 962 G; 957 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5.2e-65 Length: 4174
 Score: 1275.00 Matches: 277
 Percent Similarity: 61.42% Conservative: 94
 Best Local Similarity: 45.86% Mismatches: 141
 Query Match: 45.78% Indels: 92
 DB: Gaps: 10

US-10-023-523-44 (1-546) x ACC68992 (1-4174)

QY 3 AsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerProGlyGln 22
 DB 318 AATCACTCTGAACAGCTCTCAGCGGAACAGACAGTCAACACCT-----CCAGGTGAC 368
 QY 23 ProGluAlaGlyPro-----GluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
 DB 369 AGTTTCATCATACCCAGTCACATGGCTGGAGAGGAAGATGGCCAG-----416
 QY 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
 DB 417 -----GATTCTCCAAACCCAGTCCAAACCCAGAGAAA-----449
 QY 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
 DB 450 -----GAGGCAAGTGTGCACCCCGATATCTCTGAAGAGCTGAATCGACAGCTGGAA 500
 QY 81 AspIleLeuSerThrTyrosValAspAsnGlnGlyGlyProGlyGlyAlaAspGlyAla 100
 DB 501 GACATCATTAACACTTAT-----GGGTCTGCTGCCAGCAGCAGCGGAAGAGGCTCT 554

QY 101 GlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyTyrValAlaArgAsn 120
 DB 555 GCCAGGGCCAGTGCAGAGCTCAGAAATGCAGAAATCACCTGACAAAC-----GAGGAT 605
 QY 121 GlyGluProGluProThrProValValAsnGlyGluLysGluProSerLysGlyAspPro 140
 DB 606 GGGGACTGTGAGAAACAACCTGAAGAGCTGGAGAGAACCCGTTGCTTCTGGAGAGCA 665
 QY 141 AsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHisArgProGln 160
 DB 666 CCCACTGTCTCAAA-----GAGCCCTGAGCAATAAGGAGCAAAA-----TTG 707
 QY 161 GluLysLysLysAlaLysGlyLysGluLysGluLysThrLeuLeuMetGlnThrLeuAsn 180
 DB 708 GAAAGAAATCTCTAAAGAGATTAGGCAAGAGCAACCTGCTAAATGCAAAATCTGAAC 767
 QY 181 ThrLeuSerThrProGluGluLysLeuAlaAlaLeuCysLysLysTyTyrAlaGluLeu 200
 DB 768 AAGTTGCAACACACCGGAAGAAAGTTTGAATTTTATTCAAGAAGTATGCTGAATTCGTG 827
 QY 201 GluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeu 220
 DB 828 GATGAACATCGTACTGACCAAGAGATTAAAGCTCTCTCCAAAGAAACACAGGTCAAAAT 887
 QY 221 ValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLys 240
 DB 888 CAAAGAAAGAGGACACCTTACAGGTGAACACAGCAGAGTATCTCTCGCTCGAAGCAA 947
 QY 241 LeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGluVal 260
 DB 948 TTGAGAGTCTGTGCGGAGCTGCAGAGACACACACAGACTCTCAGAGAAAGAGCGCTT 1007
 QY 261 GlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGluValThr 280
 DB 1008 CAGCGGCACGTGAGAAAGAGAAAGAGAAATTCACAGCCATTCACAGTACC 1067
 QY 281 LeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGln 300
 DB 1068 CTCACGACATCCAGGGCCAGATCGAGCAGCAGAGTGAGCGAAATATGAAGCTCTGTGAG 1127
 QY 301 GluAsnMetGluLeuAlaGluArgLeuLysLysLeuLeuGlnTyrGluLeuArgGlu 320
 DB 1128 GAGAACACAGAGCTTGCAAAAGCTGAAAGACTCATCGATCAGTATGAGCTCAGAGAG 1187
 QY 321 GluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeuValAspAlaLys 340
 DB 1198 GAGCATCTGGACAAATATTTAAACACAGACAGACTGCAGCAGAGAGTGTGATGCAAG 1247
 QY 341 LeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGlnArgGluLysAsp 360
 DB 1248 CTTGAGCAGGCCCAAGAAATGATGAAGAGCGGAGGAGCACAACACGAGAAAGGAA 1307
 QY 361 PheLeuLysLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlu 380
 DB 1308 TATTTGCTGACACGAGCAGAGTGGAACTTCAGCGCAAGTCTCAGAGCAAGAG 1367
 QY 381 ThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThr 400
 DB 1368 ACAGTCTCTGAGGCTCAGCTCACTCTCTACTCAGAAAGTTTGAAGATTCAGAGACACA 1427
 QY 401 LeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLys 420
 DB 1428 CTAATAAAGCAACAGAGTGTTCACACGTTTCAACAGGAATGGACAAACACACTAAG 1487
 QY 421 LysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArgTrpGluSerSerAsn 440
 DB 1488 AAAATGAAGAGCTGGAAGAGCAGACACACACCTGGAAGCCCGATTTGAGACTGAAC 1547
 QY 441 LysAlaLeuLeuGluMetAlaGluLysThrValArgAspLysGluLeuGluLysLeu 460
 DB 1548 AAGGCTCTGTGGACATGATTGAAGAGAAACCATCAGAGCTGAAGAAATATGATGCTGTT 1607
 QY 461 GlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAsp 480

```

Db 1608 GTGATGAAATCGGAGGCTAGAGACCTCTCCGCTGCTTTACAGAGAGAGAAACGAA 1667
QY LeuAsnLysArgValGlnAspLeuSerAlaGlyGlnGlySerLeuThrAspSerGly 500
Db 1668 CTCACAAAAAATCAGAGACGAGAGAAATATCTGAAAGAGATGACCAAAAGTCAGCACAA 1727
QY 501 ProGluArgArgProGluGly 507
Db 1728 TCGATGAAGAGCCAGAGTCAACAGTCTCTGTGGATCAAGAGATTGACGACAGAGGTT 1787
QY 507 507
Db 1788 AATAGTGTCCAAACCGCGTGAATAATCTGGCCACAGCCTTCATGATAATTATCATCCA 1847
QY 507 507
Db 1848 GAGTCAACCCGACACAGTCCAAAGAAACCCAAACCCGAAATAGGAGTCTTCAGAGAGT 1907
QY 508 508
Db 1908 GCTGAGCGCGCTCTCAAGGAGCCAGAGCAACCCCTCTGATCCCTTCACGGGATTCAGAG 1967
QY 521 AlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGlu 540
Db 1968 AGTCCCTCGCTCCCTAACTCTCAGGCTGAAGCCGAGGAGGAGTGTCTGAACCT 2027
QY 541 ProThrSerAla 544
Db 2028 CCTCCAGGCC 2039
RESULT 24
AAX80156
ID AAX80156 standard; cdna; 4723 BP.
AC AAX80156;
XX
XX
DT 16-AUG-1999 (first entry)
DE Neurite extending activity protein encoding cdna.
KW Neurite extending activity; anti-dementia; memory; brain function;
KW dementia; ds.
XX
OS Mus sp.
FH Key
FT CDS
FT 196..2226
FT /*tag= a
XX
PN JP11147897-A.
XX
XX 02-JUN-1999.
XX
XX 13-NOV-1997; 97JP-00331242.
XX
XX 13-NOV-1997; 97JP-00331242.
XX
XX (AGEN ) AGENCY OF IND SCI & TECHNOLOGY.
XX
XX WPI; 1999-379889/32.
XX
XX DR P-PSDB; AAY17863.
XX
XX New protein with neurite extending activity - useful for treating
XX dementia.
XX
XX Claim 3; Page 5-8; 9pp; Japanese.
XX
XX The present sequence encodes a protein (I) which has neurite extending
XX activity. (I) may be administered to patients to prevent dementia or to
XX improve memory and brain function. (I) is also used as a neurite-
XX extending agent

```

```

SQ Sequence 4723 BP; 1510 A; 1038 C; 1025 G; 1150 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.48e-63 Length: 4723
Score: 1251.00 Matches: 276
Percent Similarity: 65.52% Conservative: 85
Best Local Similarity: 50.09% Mismatches: 135
Query Match: 44.92% Indels: 55
DB: 2 Gaps: 12

US-10-023-523-44 (1-546) x AAX80156 (1-4723)
QY 12 LysGlnSerAsnProLysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGln 31
Db 220 AAGCAGCAGCAGGTTTACCCTTACCTTACCGCAGACACACGAGGG 264
QY 32 GluArgProSerGlnAlaProAlaValGluAlaGluGlyProGlySerSerGlnAla 51
Db 265 CAAAGCAAGCGGAACTGTCCCGCTCTCACAGCCCTTTCCCCC---ACAAACCAAC 321
QY 52 ProArgLysProGluGlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspVal 71
Db 322 AGTCCCGAGCTGAGTGGCCACGTGT-----GATC 354
QY 72 SerGluGluSerArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsn 91
Db 355 TCAGAGAGCTGAACCGCACTGGAGGACATTATAAAACATAT----- 399
QY 92 GlnGlyGlyProGlyGluAspGlyAla-----GlnGlyGluProAlaGluPro 107
Db 400 -----GGGTCTGCAGCGAGTCTGGTAGAAGAGGAGGCACTACAGCAAACT 447
QY 108 GluAspAlaGluLysSerArgThrTyrValAlaArgAsnGlyGluProGluProThrPro 127
Db 448 GACAAGCCAGAGAGGAAGATGTGGCAGTATGGAGGATGCAGAGTGTGAGAT----- 501
QY 128 ValValAsnGlyGluLysGluProSerLysGlyAspProAsnThrGluGluIleArgGln 147
Db 502 ---GTAATGAAGAAAGTGAAGAAACCAACCACTCTCGA---GATGTTCAAGAGCA 555
QY 148 SerAspGluValGlyAspArgAspHisArgProGlnGluLysLysLysAlaLysGly 167
Db 556 AAGAGCCCGAGTCCAGCAGAGGAAACAAAG-----CTGGAGAAAGAAATCCTCAAAGGA 609
QY 168 LeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGlu 187
Db 610 CTAGGGAAGGAAGCAACCTCTCATGAAAGCTTGAACAAGCTGACTACTCCAGAGGAG 669
QY 188 LysLeuAlaAlaLeuCysLysLysTyrAlaGluLeuGluGluHisArgAsnSerGln 207
Db 670 AAGCTGACCTGTTATTAAAGATATGCTGAGTTGCTTGGAGGAGCATCTGTGAGCAG 729
QY 208 LysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeu 227
Db 730 AAGCAGCTCAAGTACCTGCAGAGAGGAGGCGCCAGATCAACAGGAGAGAGGACCACTTG 789
QY 228 ArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGlu 247
Db 790 CAGAGTCAGCAGCAGCGGACCTCTTCTGCGAGCAAGCTTGAGAGCCTGTGCGCGGAG 849
QY 248 LeuGlnArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGluGluGlu 267
Db 850 CTCAGAGGACACAAACAAACCTCAAGGAGGAAACCAATTCAGCGGCGCAGCGGAGGAGAT 909
QY 268 GluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGln 287
Db 910 GAGAAGAGGAAAGAAATAACAAATCATTTCCAGGGCAGCGCTGAGTGAATTCACGGTTCAG 969
QY 288 MetGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGluLeuAlaGlu 307
Db 970 ATTGAGCAGCAAGTGAAGAGAACATGAAGCTTGCAGGAGAGACACAGAGCTGGCAGAG 1029
QY 308 ArgLeuLysLysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAspLysValPhe 327

```


QY 136 rLysGlyAspProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspA:GAspHi 156
 Db 327 CCCTGGGGG-----GAAGCTCGAACAGATCCCTCGATGCTCAGCAAGATTC 374
 QY 156 sArgArgProGlnGluLysLysAlaLysGlyLeuGlyLysGluLysThrLeuLeuMe 176
 Db 375 AGAGTGCAACAGGAAACAAGAAACT-----TTAGCAAAAGAGTTTATTACTGAT 428
 QY 176 tGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysLys 196
 Db 429 GCAAGCCCTAACACCTTTCAACCCAGAGAGAGCTGCAGCTCTCTGTAGAAATA 488
 QY 196 rAlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuGlnLysLys 216
 Db 489 TGCTGATCTCTGGAGGAGAGCAGGAGTGTTCAGAAAGCAATGAAGATCTCGCAAGAA 548
 QY 216 sGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaVal 236
 Db 549 CCAAGCCAGATTGTGAAGAGAAAGTTCACTTCAGAGTGAACATAGCAAGGCTATCTT 608
 QY 236 uAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeu 256
 Db 609 GCGAAGCAAGCAAGTAGAATCTCTTTGCAGAGAACTTCAGCGTCACAATAAGAGCTTAA 668
 QY 256 sGluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHi 276
 Db 669 GGAGAAATAATGCAAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 728
 QY 276 sPheGlnValThrLeuAsnAspLysGlnLeuGlnMetGluGlnHisAsnGluArgAsn 296
 Db 729 TTTCCAGATTACCTTAATGAATTCAGCCAGCTGCAGCAGCATGACATCCCAACGC 788
 QY 296 rLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuGluGln 316
 Db 789 CAAACTCCGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 848
 QY 316 rGluLeuArgGluGluHisLysLysValPheLysHisLysAspLeuGlnGlnLe 336
 Db 849 CGCACTGAGGAGAGAGACATTTATAGGTGTTCAACATAGCAATGCGACAGCT 908
 QY 336 uAlaAspAlaLysLeuGlnGlnGlnGlnMetLeuLysGluAlaGluArgHisG 356
 Db 909 CGTGGATGCCAACTGCAGCAACAGCACACATGATAAAGAGCTGATCAAAACATCA 968
 QY 356 nArgGluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeu 376
 Db 969 GAGAGAGAGAGGTTTTTTTAAAGAGCGAGAGATCGAGGCAAAATAGCAAAAT 1028
 QY 376 tLysGlnGlnGlnThrHisLeuLysGlnGlnLeuAlaLeuTyThrGluLysPheGlu 396
 Db 1029 GAAACAGCAAGAGTACAACTAAACAGCAGCTTTCTTTTATGATAGATTGAGA 1088
 QY 396 uPheGlnAsnThrLeuSerLysSerSerGluValPheThrPheLysGlnGluMet 416
 Db 1089 ATTCCAGACTACCTGCGCAAAAGCAATGAACTGTTTACAACTTCAGACAGGAAATGGA 1148
 QY 416 uLysMetThrLysLysLysLysLysLeuGluLysGlu 428
 Db 1149 AAAGATGACAAAGAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1185
 RESULT 27
 AAF22469
 ID AAF22469 standard; cDNA; 823 BP.
 XX
 AC AAF22469;
 XX
 DT 26-MAR-2001 (first entry)
 XX
 DE Human breast cancer associated antigen nucleotide sequence SEQ ID NO:48.
 XX
 KW Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
 KW cancer associated antigen; cytostatic; cancer vaccine; ss.

XX Homo sapiens.
 OS
 XX
 PN W0200073801-A2.
 XX
 PD 07-DEC-2000.
 XX
 PF 26-MAY-2000; 2000MO-US014749.
 XX
 PR 28-MAY-1999; 99US-0136526P.
 DR 10-SEP-1999; 99US-0153454P.
 XX
 PA (LUDW-) LUDWIG INST CANCER RES.
 XX
 XX Obata Y;
 XX
 XX WPI; 2001-025274/03.
 DR
 XX Nucleic acids encoding breast, gastric and prostate cancer associated
 PT antigen precursors, useful for diagnosing and treating a condition
 PT characterized by expression of an abnormal amount of a protein, e.g.
 XX cancer.
 XX
 PS Claim 50; Page 289-290; 799pp; English.
 XX
 CC AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
 CC represent nucleotide sequences encoding human breast, gastric and
 CC prostate cancer associated antigen precursors (CAAP) respectively.
 CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
 CC represent human breast, gastric and prostate CAAP protein sequence
 CC respectively. CAAPs have cytostatic activity and can be used in the
 CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
 CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
 CC condition characterised by expression of an abnormal amount of a protein,
 CC e.g. cancer
 XX
 SQ Sequence 823 BP; 324 A; 155 C; 192 G; 149 T; 0 U; 3 Other;
 Alignment Scores:
 Pred. No.: 1.98e-45 Length: 823
 Score: 925.50 Matches: 197
 Percent Similarity: 86.55% Conservative: 41
 Best Local Similarity: 71.64% Mismatches: 37
 Query Match: 33.23% Indels: 5
 DB: Gaps: 0
 US-10-023-523-44 (1-546) x AAF22469 (1-823)
 QY 179 LeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysLysLys 198
 Db 2 CTAACACCTTTCAACCCAGAGAGAGCTGCAGCTCTCTGTAGAAATATGCTGAT 61
 QY 199 LeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuGlnLysLysGlnSer 218
 Db 62 CTTCTGGAGGAGAGCAGGAGTGTTCAGAAAGCAATGAAGATCTCGCAAGAGCAAGCC 121
 QY 219 GlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArg 238
 Db 122 CAGATTGTGAAGAGAAAGTTCACTTCAGAGTGAACATAGCAAGGCTATCTTGGCAAGA 181
 QY 239 SerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGlu 258
 Db 182 AGCAAGCTAGAAATCTTTTCAGAGAACTTCAGCGTCACAATAAGACGTTAAGAGAGGA 241
 QY 259 GlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGln 278
 Db 242 AATATGCAAGGAGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 301
 QY 279 ValThrLeuAsnAspLysGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeu 298
 Db 302 ATTACCTTAAATGAATTCAGCCAGCTGCAGCAGCATGACATCCCAACGCAACTC 361
 QY 299 ArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLysLysLysLysGluLeu 318

Db	362	CGACAGGAAAACATTCAGCTCGGGGAGAAAGCTCAAGAAAGCTCATCGAACAGTACGCACGTG	421
QY	319	AzqGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeuValAsp	338
Db	422	AGGAGAGGACCATTCATGATAGGTGTTCAACATPAGGAACCTGCAACAGCAGCTCGTGGAT	481
QY	339	AlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluGluArgHisGlnArgGlu	358
Db	482	GCCAACTGCAGCAAAACGACACCACTGATAAAGAGCTGATGAAAAACATCAGAGAGAG	541
QY	359	LysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGln	378
Db	542	AGAGAGTTTTTTATTAAGAAGACGACAGATCGAGGCACAAATACGAACAATGAACAG	601
QY	379	GlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGln	398
Db	602	CAAGAAGTACACTA-AAACAGCAGCTTCTCTTTATATGATAAGTTTGAAGAATTTCAG	660
QY	399	AnnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMet	418
Db	661	ACTACATCGGCACAAAAGCAATGAAGTGTTCACACCTTCAGACAGGAATGGGAAAAGAT	719
QY	419	ThrLysLysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArgTrpGluSer	438
Db	720	GCCAGAAAATTAA-AAACTGGGAAAGAAACCATTAATTGGGGTTCCTCCATGGGAAAC	778
QY	439	SerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArg	453
Db	779	AATATTAAACN-CTTTTGCA-ATGGNTGAAGAAAAACAGTCCGG	821
RESULT 28			
AAI88232			
ID	AAI88232 standard; cDNA; 1501 BP.		
XX	AAI88232;		
XX	06-NOV-2001 (first entry)		
DE	Human polynucleotide SEQ ID NO 8292.		
XX	Human; cytokine; cell proliferation; cell differentiation; gene therapy;		
KW	vaccine; peptide therapy; stem cell growth factor; haematopoiesis;		
KW	tissue growth factor; immunomodulatory; cancer; leukaemia;		
XX	nervous system disorders; arthritis; inflammation; ss.		
XX	Homo sapiens.		
XX	WO200164835-A2.		
XX	07-SEP-2001.		
XX	26-FEB-2001; 2001WO-004927.		
XX	28-FEB-2000; 2000US-00515126.		
XX	18-MAY-2000; 2000US-00577409.		
XX	(HYSE-) HYSEQ INC.		
XX	Tang YT, Liu C, Drmanac RT;		
XX	WPI; 2001-514838/56.		
XX	P-PSDB; AAC08301.		
XX	Isolated nucleic acids and polypeptides, useful for preventing diagnosing		
PT	and treating e.g. leukemia, inflammation and immune disorders.		
XX	Claim 1; SEQ ID NO 8292; 1399pp + Sequence Listing; English.		
XX	The invention relates to human polynucleotides (AAI79941-AAI93841) and		
CC	the encoded proteins (AAC00010-AAC013910) that exhibit activity elating to		
CC	cytokine, cell proliferation or cell differentiation or which may induce		
CC	production of other cytokines in other cell populations. The		

CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pat_sequences
 XX
 SQ Sequence 1501 BP: 473 A; 375 C; 393 G; 259 T; 0 U; 1 Other;
 XX

Alignment Scores:	
Pred. No.:	1.96e-38
Score:	810.00
Percent Similarity:	63.22%
Best Local Similarity:	47.70%
Query Match:	29.08%
DB:	Indels: 60
	Gaps: 2
	4

	Length:	1501
	Matches:	166
	Conservative:	54
	Mismatches:	68
	Indels:	60
	Gaps:	2
		4

US-10-023-523-44 (1-546) x AA198232 (1-1501)

Qy	257	GluGluGlyValGlnArgAlaArgGluGluGluGluLysArgLysGluValThrSerHis	276
Db	87	GAAGAGGCGCTTCAGCGGCGACGTCGAGAAAGAGAAAGAGAAATCAAGCCAT	146
Qy	277	PheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSer	296
Db	147	TTCCAGAGTACCTTCAGGACATCCAGGGCCAGATCGAGCAGAGTAGTGAGCGAAATATG	206
Qy	297	LysLeuArgGlnGlnAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTyr	316
Db	207	AAGCTCTCTCAGGAGAACACAGAGCTTGCAAAAAGCTGAAAAGCATCATCGATCASTAT	266
Qy	317	GluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeu	336
Db	267	GAGCTCAGAGGAGGAGCATCTGACAAAATATTTAAACACAGAACTGCAGCAGAAAGCTG	326
Qy	337	ValAspAlaLysLeuGlnGlnAlaGlnGlnMetLeuLysGluAlaGluArgHisGln	356
Db	327	GTGGATGCAAAAGCTTGAGCAGCGCCCAAGAAATGATGAAGAGAACGGAGGAGCGACAA	386
Qy	357	ArgGluLysAspPheLeuLysGluAlaValGlnLysGlnArgMetCysGluLeuMet	376
Db	387	CAGAAAAGGAATATTTCCTGAACGAGCAGCAGTGGAAACTTCAGCGGAAAGTGCTG	446
Qy	377	LysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGlu	396
Db	447	AAGGACGAGAGACAGATCCTCCAGGCTCAGCTCACTCTCTACTCAGAAAGTTTGAAGAA	506
Qy	397	PheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGlu	416
Db	507	TTCCAGACACACTAACTTAAAGAACGACGAGTGTTGCCACGTTCTAAACAGGAAATGGAC	566
Qy	417	LysMetThrLysLysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArgTyr	436
Db	567	AAAACACTAGAAAATAGAGAGCTGAAAAGACACACCCACATGGAAAGCCCATTT	626
Qy	437	GluSerSerAsnLysAlaLeuLeuGluMetalGluLysThrValArgAspLysGlu	456
Db	627	GAGAACTGTAAACAAGCTCTCTTGCACATGATTGAAGAGAAAGCAGCTAGAGAGCTAAAGAA	686
Qy	457	LeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThr	476
Db	687	TATGATGCTTGTGATGAAATCGGAGGCGTAGAAGCACTCTGCCCTGCTTTTACAGAA	746
Qy	477	GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeu	496
Db	747	GAGAAACAGCACTCCACAAAATAATCAGACCCAGAAATATCTGAAAAGGATGACCAA	806
Qy	497	ThrAspSerGlyProGluArgProGluGly	507
Db	807	AGTCAGCACTCCGATGAAGACGACGAGTCAAACCTCTCTGTGATCAAGAGATTGAC	866

expression; (c) comparing the results with a control; and (d) determining whether the comparison indicates a likelihood disease). (M1) is useful for identifying modulators of the p53 pathway. A probe for HM expression is useful for diagnosing breast, colon, kidney, lung and ovarian cancer, in a patient, where the cancer has greater than 25 % expression level. Modulators identified by (M1) are useful in a variety of diagnostic and therapeutic applications, where disease or disorder prognosis is related to defects in the p53 pathway, such as, angiogenesis, apoptotic or cell proliferation disorders (e.g. cancer). Another two new methods (M2 and M3) are useful for modulating the p53 pathway of a cell, thus restoring the p53 function of the cell, so that the cell undergoes normal proliferation or progression through the cell cycle. (M2) and (M3) are also useful for treating defects in the p53 pathway such as angiogenic, apoptotic or cell proliferation disorders. The present sequence is an HM nucleic acid encoding a p53 pathway modifying protein.

```

OY 58 aGlnAlaThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgG1 78
DB 70574 TCAAGCCAGAACGGCTCAGTCTGGGGCCCTTCGTGATGTCCTCAGGAGCTGAGCCGCCA 70633
OY 78 nLeuGluAspIleLeuSerThrTyrcysValAspAsnAsnGlnGlyGlyProGlyGluAs 98
DB 70634 ACTGGAACACATACTGACACATATCTGTGTGACATAAACCCAGGGGGCCCGCGAGGA 70693
OY 98 pGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyValAl 118
DB 70694 TGGGGCACAGGGTGGCGGCTGAACCCGAAGATCGAGAGAATGCCCGGACCTATGTGGC 70753
OY 118 aArgAsnGlyGluProGluProThrProValValAsnGlyGluLysGluProSerLysG1 138
DB 70754 AAGGATGGGAGCTGAACTCACTCCAGTAGTCAATGGAGAGAGAACCTCCAGGG 70813
OY 138 yAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHisArgAr 158
DB 70814 GGATCCAAACACAGAGATCCGGCAGAGTCCGAGGTCGGAGACCGAGACCATCGAAG 70873
OY 158 gProGlnGluLysLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGlnTh 178
DB 70874 GCCACAGGAGAGAAAAAAGCCAGGGTTTGGGTGAGCAGAGGGCGGCTCTTTGTGAA-- 70931
OY 178 rLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuLysLysLysTyAlaG1 198
DB 70932 -----GCTGGTGAGGAGAGGAGCTTTGGACTTCGACCTTCTCTGGGCGAG 70975
OY 198 uLeuLeu-----GluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuG1 214
DB 70976 TCTGTCTGCCAGGATTCAAAGGAAACGGTACTTCTCAGAGCAGC---AAGTCACTCTA 71032
OY 214 nLysLysGlnSerGln 219
DB 71033 GTCTAATCAAGCCAG 71048
RESULT 33
AAH26497
ID AAH26497 standard; DNA; 22255 BP.
AC AAH26497;
XX
XX 12-NOV-2001 (first entry)
XX
XX Human low density lipoprotein binding protein 3 (LBP-3) gene.
DE
DE Low density lipoprotein binding protein 3; LBP-3; LDL; human;
KW atherosclerosis; antiarteriosclerotic; gene therapy; diagnosis; vaccine;
KW ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH 2100..16932
CDS /tag= a
FT /note= "contains introns"
FT 2100..2288
FT /tag= b
FT 2269..2296
FT /tag= c
FT 2297..3332
FT /tag= d
FT 3333..6309
FT /tag= e
FT 6310..6401
FT /tag= f
FT 6402..9738
FT /tag= g
FT 9739..9909
FT /tag= h
FT 9910..11839
FT /tag= i
FT 11840..12034

```

```

FT intron /tag= j
FT 12035..14095
FT /tag= k
FT 14096..14215
FT /tag= l
FT 14216..14443
FT /tag= m
FT 14444..14518
FT /tag= n
FT 14519..14944
FT /tag= o
FT 14945..15037
FT /tag= p
FT 15038..15801
FT /tag= q
FT 15802..15897
FT /tag= r
FT 15898..16638
FT /tag= s
FT 16639..16932
FT /tag= t
DN WO200164874-A2.
XX
XX 07-SEP-2001.
XX
XX 28-FEB-2001; 2001WO-US006356.
XX
XX 02-MAR-2000; 2000US-00517849.
XX 14-JUL-2000; 2000US-00616289.
XX
XX (BOST-) BOSTON HEART FOUND INC.
XX
XX Lees AM, Lees RS, Law SW, Arjona AA;
XX WPI; 2001-565505/63.
XX P-FSDB; AAB82809.
XX
XX New isolated low density lipoprotein binding polypeptide for treating,
XX diagnosing and/or identifying therapeutic agents for atherosclerosis.
XX
XX Example 4; Fig 24; 143pp; English.
XX
XX The present sequence is that of genomic DNA encoding novel human low
XX density lipoprotein binding protein 3 (LBP-3, see AAB82809). The DNA was
XX isolated from a human genomic library by screening with LBP-3 cDNA (see
XX AAH26501). The open reading frame spans 10 exons. Human LBP-2 nucleic
XX acids are among claimed polynucleotides of the invention that encode
XX novel polypeptides, termed LBPs, capable of binding to native and
XX methylated LDL. Also claimed are isolated LBP polypeptides, and
XX biologically active fragments and analogues of them, as well as
XX expression vectors, cells and methods of producing the LBPs. Methods of
XX determining if an animal is at risk for atherosclerosis, methods for
XX evaluating an agent for use in treating atherosclerosis, and methods for
XX treating a cell having an abnormality in structure or metabolism of LBP
XX are claimed. Pharmaceutical compositions comprising an LBP polypeptide or
XX nucleic acid, and vaccine compositions, are also claimed
XX
XX Sequence 22255 BP; 5195 A; 5302 C; 5924 G; 5834 T; 0 U; 0 Other;
SQ

```

Alignment Scores:

```

Pred. No.: 3,78e-33 Length: 22255
Score: 740.50 Matches: 182
Percent Similarity: 40.56% Conservative: 7
Best Local Similarity: 39.08% Mismatches: 23
Query Match: 26.59% Indels: 255
DB: 5 Gaps: 4

```

US-10-023-523-44 (1-546) x AAH26497 (1-22255)

```

OY 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerPro 20
DB 2100 ATGAAGAACCAAGACAAAAGAACGGGGCTGCCAACCAATCAATCCAAAAGACGCCCA 2159

```


QY 138 GlyAspProAsnThrGluGluIleuLeuArgGlnSerAspGluValGlyAspArgAspHisArg 157
Db 3239 GGGGATCCAAACACAGAGAAGATCCCGCAGAGTGACGAGGTGCGAGACCGAGACCATCGA 3298
QY 158 ArgProGlnGluLysLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGln 177
Db 3299 AGGCCACAGAGGAAGAAAAAAGCCAGGGTTTGGGTGAGCAGAGGGCGGCTCTTTGTGAA 3358
QY 178 ThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysTyrAla 197
Db 3359 -----GCTGCTGAGCAGAGGAGGTTCGACTTGACGTTCTCTGGGCC 3400
QY 198 GluLeuLeu-----GluGluHisArgAsnSerGlnLysGlnMetLysLeuLeu 213
Db 3401 AGTCTGTTGCGCAGGATTCAAGGAANAACGTACTTCTCAGACAGC---AAGTCACCTC 3455
QY 214 GlnLysLysGlnSerGln 219
Db 3458 TAGCTTAATCAAGCCAG 3475
RESULT 34
AAS66673
ID ID AAS66673 standard; cDNA; 1233 BP.
XX AC AAS66673;
XX XX
DT 13-FEB-2002 (first entry)
XX XX
DE DNA encoding novel human diagnostic protein #2477.
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX PD 11-OCT-2001.
XX XX
XX 30-MAR-2001; 2001WO-US008631.
XX XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX XX
PA (HYSE-) HYSEQ INC.
XX XX
PI Drmanac RT, Liu C, Tang YT;
XX XX
XX WPI; 2001-639362/73.
DR P-PSDB; AEG02486.
PT
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 2477; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (II) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pat_sequences

XX SQ Sequence 1233 BP; 438 A; 230 C; 287 G; 278 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 8, 9e-29 Length: 1233
 Score: 64250 Matches: 153
 Percent Similarity: 55.61% Conservative: 55
 Best Local Similarity: 40.91% Mismatches: 85
 Query Match: 23.07% Indels: 81
 DB: 5 Gaps: 11

US-10-023-523-44 (1-546) x AAS66673 (1-1233)

QY 183 SerThrProGluGluLysLeuAlaLeuCys-----LysLysTyrAlaGluLeuLeu 200
 DB 253 AGCAATGGAAGAA-----GCTGGACTTTGTGGTTAGAGAGAAAGCAGATATGTTG 306
 QY 201 -----GluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGln 217
 DB 307 TGTAACCTGTAATCAAT-----GATATCTTCAACATCAAGAC 345
 QY 218 Ser-----GlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaVal 235
 DB 346 TCAAAATTCAGTGCACCAAGTAATAACATTATTGGAAGATGAAGAGCCGCGACTTT 405
 QY 236 LeuAlaArgSerLysLeuGluSerLysLeuGluGlnArgHisAsnArgSerLeu 255
 DB 406 ATACAAAG-----AACAGAGGTGG 426
 QY 256 LysGluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSer 275
 DB 427 ---GAGGAACATATGCAGCAAAAAGAGAGAGAGAGAGTCTTAAAGAGTAACGTGA 483
 QY 276 HisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsn 295
 DB 484 CATTTCCAAATTAATTTAACTGAACACTCAAGCCAGCTGGACAGCATGAATATACACAAT 543
 QY 296 SerLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLysLysLeuIleGluGln 315
 DB 544 GCCAACTGCAGAGAGACATCGAATGGAGAGAAAGCTTAAGAAAGCTCACTGACCAG 603
 QY 316 TyrGluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGln 335
 DB 604 TATGCACTGAGGAAGAGCAAAATTAATAAGCGTTCAACATAAGGAATATGACCAACAG 663
 QY 336 LeuValAspAlaLysLeuGlnGlnAlaGlnLysMetLysLysGluAlaGluArgHis 355
 DB 664 CTGTAGATGCCAGACTCAACAAACAGCAGCTGATAAAGAGAGCTGATGAAGACAT 723
 QY 356 GlnArgGluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeu 375
 DB 724 CAGAGAGAGAGAGATTTTATTAAAGAGAGCAACAGATCGAGCAACAATATGAAGAA 783
 QY 376 MetLysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuThrThrGluLysPheGlu 395
 DB 784 ATGAACAAGAGAGAGAGCAACACTGAAGAGAGAGCTTTCTTTATATGGATAGTTGAA 843
 QY 396 GluPheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMet 415
 DB 844 GAATTCAGACTACCATGGCAAAACCAATGAATTTTACAGCTTCAAGCAGGAACG 903
 QY 416 GluLysMetThrLysLysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArg 435
 DB 904 GAAAG-----
 QY 436 TrpGluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLys 455

DB 910 -----AAACTATTCTGATATAA 927
 QY 456 GluLeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGln 475
 DB 928 AATTATAGGCTTTCAATAAAATTTGGAGCGTTAGAGAGCTGTACAAAGGCTTTCAA 987
 QY 476 ThrGluArgAsnAspLeuAsnLysArgVal----- 485
 DB 988 ATAGAAAGGAATGAATCAGTCAGTCAGAAACTGGGAATTTGAAAGGCGAGCTCTCTGTGAAA 1047
 QY 486 -----GlnAspLeuSerAlaGlyGlnGlySerLeuThrAspSerGlyProGlu 502
 DB 1048 GTAGCAGATGTAGATTAGCAGCTGCTGTGAGCAATCTCTGTGCTGACCTGGATTTTCC 1107
 QY 503 Arg-----ArgProGluGlyProGlyValAlaGlnAlaProSerSerProArg 517
 DB 1108 AATATGCTGAACACTTCTCTTAAAGAGCCCGAGAGTCCATCTGGAGGCTGACCCCAA 1167
 RESULT 35
 AAS92278/C
 ID AAS92278 standard; cDNA; 1734 BP.
 AC AAS92278;
 DT 13-FEB-2002 (first entry)
 XX DNA encoding novel human diagnostic protein #28082.
 DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX Homo sapiens.
 XX WO200175067-A2.
 XX 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US008631.
 XX 31-MAR-2000; 2000US-00540217.
 XX 23-AUG-2000; 2000US-00649167.
 XX (HYSE-) HYSEQ INC.
 XX Drmanac RT, Liu C, Tang YT;
 XX WPI; 2001-639362/73.
 XX P-PSDB; ABG28091.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 XX diagnostics, forensics, gene mapping, identification of mutations
 XX responsible for genetic disorders or other traits and to assess
 XX biodiversity.
 XX Claim 1; SEQ ID NO 28082; 103pp; English.
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 XX sequences. (I) is useful as hybridisation probes, polymerase chain
 XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 XX and in recombinant production of (II). The polynucleotides are also used
 XX in diagnostics as expressed sequence tags for identifying expressed
 XX genes. (I) is useful in gene therapy techniques to restore normal
 XX activity of (II) or to treat disease states involving (II). (II) is
 XX useful for generating antibodies against it, detecting or quantitating a
 XX polypeptide in tissue, as molecular weight markers and as a food
 XX supplement. (II) and its binding partners are useful in medical imaging
 XX of sites expressing (II). (I) and (II) are useful for treating disorders
 XX involving aberrant protein expression or biological activity. The
 XX polypeptide and polynucleotide sequences have applications in

Sequence 1734 BP; 484 A; 404 C; 327 G; 519 T; 0 U; 0 Other;

US-10-023-523-44 (1-546) x AAS92278 (1-1734)

SQ Sequence 1843 BP; 591 A; 424

Sequence 1843 BP; 591 A; 424 C; 508 G; 320 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4.52e-23 Length: 2526
Score: 550.00 Matches: 121
Percent Similarity: 66.53% Conservatives: 40
Best Local Similarity: 50.00% Mismatches: 51
Query Match: 19.75% Indels: 30
DB: 5 Gaps: 7

US-10-023-523-44 (1-546) x AAS66674 (1-2526)

QY 183 SerThrProGluGluLysLeuAlaLeuLysCys-----LysLysTyrAlaGluLeuLeu 200
DB 157 AGCAATGGAAGAA-----GCTGGACTTGTGGTTAAGAGAGAAAGCAGATATGTTG 210
QY 201 -----GluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuLysLysGln 217
DB 211 TGTAACTGTAATCAGAT-----GATATCTTCAACATCAAGAC 249
QY 218 Ser-----GlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaVal 235
DB 250 TCAAAATTCGATGCCACAGTAATAAACATTTATTGGAAGATGAAGAGCCGCTGATTT 309
QY 236 LeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeu 255
DB 310 ATACAAAG-----AACAGAGTTGG 330
QY 256 LysGluGluValGlnArgAlaArgGluGluGluLysArgLysGluValThrSer 275
DB 331 --GAGGAACATATGTCAGCAGCAAAAAGAGAGAGAGAGAGTCTTAAAGAAAGTAACATGCA 387
QY 276 HisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsn 295
DB 388 CATTTCCAAATTAATCTTAACTGAACTCAAGCCAGCTGGAACAGCATGAAATACACAAT 447
QY 296 SerLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGluGln 315
DB 448 GCCAACTGCAGCAGGAGAGACATGAAATGGAGAGAAAGCTAAAGAGCTCACTGACCAG 507
QY 316 TyrGluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGln 335
DB 508 TATCATCTGGAGGAGAGCAAAATTAATAAGCGTTCAACATATAGGAATACGACACAG 567
QY 336 LeuValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluGluArgHis 355
DB 568 CTGTGATGCGCACTTCAACAAACAGCAGCAGCTGATAAAGAGCTGTATGAAGACAT 627
QY 356 GlnArgGluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeu 375
DB 628 CAGAGAGAGAGAGAGTATTATTAAGAGAGCAACAGATCGAGGCACAAATATGAAGA 687
QY 376 MetLysGlnGlnThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGlu 395
DB 688 ATGAACAAGAGAGAGCAAACTGAAGAGAGCGCTTTTCCTTTATATGATAAGTTTGA 747
QY 396 GluPheGlnAsnThrLeuSerLysSerGluValPheThrThrPheLysGlnGluMet 415
DB 748 GAATTCAGACTACCATCGGCAAAACCAATGAACCTTTTACAGCCTTCAAGCAGGAACG 807
QY 416 GluLys 417
DB 808 GAAAG 813

RESULT 38

ID ABL10154 standard; cDNA; 4307 BP.

XX ABL10154;

AC ABL10154;

XX 26-MAR-2002 (first entry)

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 24944.

XX Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ss.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-VAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX P-PSDB; ABB66051.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

XX genes from Drosophila and for elucidating cell signalling and cell-cell

XX interactions.

XX Claim 1; SEQ ID NO 24944; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

XX capable of detecting 1000 or more genes from Drosophila. The invention is

XX useful in developmental biology and in elucidating cell signalling and

XX cell-cell interactions in higher eukaryotes for the development of

XX insecticides, therapeutics and pharmaceutical drugs. The invention

XX discloses genomic DNA sequences (AB161617-AB130511), expressed DNA

XX sequences (AB101840-AB16175) and the encoded proteins (ABBS5737-

XX ABB2072). The sequence data for this patent did not form part of the

XX printed specification, but was obtained in electronic format directly

XX from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 4307 BP; 1298 A; 971 C; 1016 G; 1022 T; 0 U; 0 Other;

XX Alignment Scores:

XX Pred. No.: 2.59e-21 Length: 4307

XX Score: 524.00 Matches: 153

XX Percent Similarity: 47.36% Conservatives: 98

XX Best Local Similarity: 28.87% Mismatches: 115

XX Query Match: 18.82% Indels: 166

XX DB: 4 Gaps: 11

XX US-10-023-523-44 (1-546) x ABL10154 (1-4307)

QY 88 ValAspAsnAsnGlnGlyGlyProGlyGlu---AspGlyValaGlnGlyGluProAlaGlu 106

DB 1044 ATAGACCATGGAGAGCTAAGCAAGCGAAGAAAGTGGCGCGAGGAGCAAGCGCA 1103

QY 107 ProGluAspAla-GluLysSerArgThrTyrValAlaArgAsnGlyGlu---ProGluPr 125

DB 1104 CCAGAGTGGAGAGTGGTAAGA-----GCGGTGATGGCCAAAGCAAGCAAGCC 1154

QY 125 oThrProValValAsnGlyGluLysGluProSerLys-GlyAspProAsnThrGluLui 145

DB 1155 AAACCTATTTCATAGACTGCCAGATCCATGGAACCGGAATAAATAGACA----- 1207

QY 145 leArgGlnSerAspGluValGlyAspArgAspHisArgProGlnGluLysLysLys 165

DB 1208 -----AAATTTAAGG 1217

QY 165 lalysGlyLeuGlyLysGluIleThrLeu-----LeuMetGlnThrLeuAsnThrLeu 183

DB 1218 CG-----GAATTATATCTATTATTGTCATGAATCCCTTGACGAATGCC 1262

QY 183 exThr-ProGluGluLysLeuAlaAlaLeuCysLysLysTyrAlaGluLeuLeuGluH 203

DB 1263 CCTCCGCCGAGGAGAGGTGAACCT-----CTGCTGCAAGCC 1301

QY 203 iArgAsnSerGlnLys-----GlnMetLysLeuLeuGlnLysL 216

Db 2380 CTAATGCCAATGGTATCGATTTTAGCCACAGAAAGAGTCTGCAGACGCACGACTCTAGAGC 2433
 Qy 459 lyleuGlnVallyslleGlnArgLeuGlulysLeuCysArgAlaLeuGlnThrGluArgA 479
 Db 2440 GCCTGCAGAGCAACAATACAGACGCTGCAGAGCTACTCGTCTCTTACAGCTGGAGGGA 2499
 Qy 479 snAspLeuAnsnLysArgValGlnAsp 487
 Db 2500 CAACACTGCACAAAGTGCCTGCGGGAT 2525
 RESULT 39
 ABN74519
 ID ABN74519 standard; cDNA; 937 BP.
 AC AC
 XX AC
 XX AC
 XX
 DT DT
 XX DT
 XX DT
 DE DT
 DE DE
 XX DE
 KW KW
 KW KW
 XX KW
 CS CS
 CS Bos taurus.
 PN WO200194550-A2.
 XX
 XX 13-DEC-2001.
 PD
 XX
 XX 07-JUN-2001; 2001WO-US018576.
 XX
 PR 07-JUN-2000; 2000US-0209874P.
 PR 06-JUN-2001; 2001US-00876143.
 XX
 XX (INFI-) INFIGEN INC.
 PA
 XX
 XX Bilertsen KJ, Pfister-Genskow M, Childs L;
 XX WPI; 2002-351289/38.
 XX
 DR
 XX
 PT An expressed sequence tag (EST), the expression of which, or its
 PT complementary sequence, in a cell identifies the cell as a
 PT developmentally competent or incompetent cell.
 XX
 PS Example 16; Page 461-462; 584pp; English.
 XX
 CC The present invention describes an expressed sequence tag (EST), where
 CC the EST is an isolated, enriched, or purified nucleic acid sequence
 CC representing all or part of a gene, the expression of which, or its
 CC complementary sequence, in a cell identifies the cell as a
 CC developmentally competent or incompetent cell. Molecules which induce
 CC developmental competence in a cell line are useful for inducing
 CC totipotency in one or more cells. Molecules which induce developmental
 CC competence in a cell line are useful for preventing a full term
 CC pregnancy in an animal and inhibiting totipotency. The molecules are also
 CC useful for treating a disease in an animal by inducing development of one
 CC or more cells of the animal into a specific cell type. The present
 CC sequence represents a bovine EST which is given in the exemplification of
 CC the present invention
 XX
 SQ Sequence 937 BP; 222 A; 211 C; 265 G; 203 T; 0 U; 36 Other;

Db 22 GGTGGCGCCCTCTAGACTAGTGGATCCCC---GGGTCTCAGGAATT-CGGCACGAG 77
Qy 65 SerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGluAspIleLeuSer 84
Db 78 GCTGGAGCCCTCCGTGATGCTCTGAGGAGTGTGAGCCGCGAGCTGGAGACATCTCAGT 137
Qy 85 ThrTyrCysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAlaGlnGlyGluPro 104
Db 138 ACATCTCGTGGCAACAGCTCAGGGGGCCCGAGTGGAGTGGCCACAGGGTGGACCT 197
Qy 105 AlaGluProGluAspAlaGluLysSerArgThrTyrValAlaAlaArgAsnGlyGluProGlu 124
Db 198 GCTGAACCCGAAGATCGACAGAAAGTCCCGGACCTATGCTCAAGGAATGGGGACCTGGAG 257
Qy 125 Pro---ThrProValValAsnGlyGlyLysGluProSerLysGlyAspProAsnThrGlu 143
Db 258 CCAGAGACTCCAGTAGTCAATGTTGAGAGGAATCTCCAGGGGGAGCCGGCCGGAC 317
Qy 144 GluIleArgGlnSerAspGluValGlyAspArgAspHisArgProGlnGlyLysLys 163
Db 318 GAGATCCGGACCCAGTGTAGTCTGAGACCGGACCCAGGACCCACAGGANAAGAAA 377
Qy 164 LysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThr 181
Db 378 AAAGCCAAAGGTCTCGGAAAGANATCACC-----TGCCCCACTTTAACTACC 425

RESULT 40
ACH31687/c
ID ACH31687 standard; cDNA; 334 BP.
AC ACH31687;
XX
XX
XX 13-OCT-2003 (first entry)
XX
XX Human bone marrow cDNA #978.
XX
XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
XX Genome mapping; biodiversity; genetic disorder.
XX
XX Homo sapiens.
XX
XX US2003073623-A1.
XX
XX 17-APR-2003.
XX
XX 30-JUL-2001; 2001US-00918995.
XX
XX 30-JUL-2001; 2001US-00918995.
XX
XX (DRMA/) DRMANAC R T.
XX (LABA/) LABAT I.
XX (STAC/) STACHE-CRAIN B.
XX (DICK/) DICKSON M C.
XX (JONE/) JONES L W.
XX
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX WPI; 2003-615964/58.
XX
XX New polynucleotide sequences obtained from various cDNA libraries, useful
XX as hybridization probes, as oligomers for PCR, for chromosome and gene
XX mapping, in the recombinant production of protein, or in generating
XX antisense DNA or RNA.

XX Claim 1; SEQ ID NO 18899; 44pp; English.
XX
XX The invention relates to an isolated polynucleotide comprising any one of
XX 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
XX determined by the technique of SBH (sequencing by hybridisation). Also
XX included is a purified polypeptide comprising a sequence corresponding to
XX a reading frame of the novel polynucleotide. The nucleic acid sequences
XX are useful in diagnostics as expressed sequence tags (EST) for

CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversity, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030073623

XX
SQ Sequence 334 BP; 55 A; 82 C; 73 G; 124 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5,31e-19 Length: 334
Score: 464.00 Matches: 87
Percent Similarity: 91.89% Conservative: 15
Best Local Similarity: 78.38% Mismatches: 9
Query Match: 16.66% Indels: 0
DB: 8 Gaps: 0

US-10-023-523-44 (1-546) x ACH31687 (1-334)

Qy 226 HisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCys 245
Db 334 CACTTCAGAGTGAACATAGCAGAGGCTATCTTGGCAAGAGCAAGCTAGAAATCTCTTTC 275
Qy 246 ArgGluLeuGlnArgHisAsnArgSerLeuLysGluGlyValGlnArgAlaArgGlu 265
Db 274 AGAGAACTTCAGGCTCACATAAAGAGCTTAAGAGAGAAATATGCAGCAGGACGAGAG 215
Qy 266 GluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGln 285
Db 214 GAGAGAGACGAGCTAAGAGAGCACTGCACATTCAGATTACCTTAATGAATTCAA 155
Qy 286 LeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGluLeu 305
Db 154 GCCCAGCTGGAGCAGCATGACATCCACACGCCAAACTCCGACAGGAAAAACATTGAGCTG 95
Qy 306 AlaGluArgLeuLysLysLeuLysGluGlnThrGluLeuArgGluGluHisIleAspLys 325
Db 94 GGGGAGAGCTTAAGAGAGCTCATCGAACATGACATGAGGAGAGACATGATGATAG 35
Qy 326 ValPheLysHisLysAspLeuGlnGlnGlnLeu 336
Db 34 GTGTTACACATATAGGACCTGCAACAGCAGCTC 2

RESULT 41
AAH07907
ID AAH07907 standard; cDNA; 710 BP.
XX
XX
XX AC AAH07907;
XX
XX 26-JUN-2001 (first entry)
XX
XX Human cDNA clone (5'-primer) SEQ ID NO:4742.
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX Homo sapiens.
XX
XX EP1074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-00116126.
XX
XX 29-JUL-1999; 99JP-00248036.
XX 27-AUG-1999; 99JP-00300253.
XX 11-JAN-2000; 2000JP-00118776.
XX 02-MAY-2000; 2000JP-00183767.
XX 09-JUN-2000; 2000JP-00241899.

XX PA (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

XX Claim 1; SEQ ID NO 4742; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention

XX SQ Sequence 710 BP; 224 A; 144 C; 203 G; 136 T; 0 U; 3 Other;

Alignment Scores:

Pred. No.:	4,54e-13	Length:	710
Score:	368.00	Matches:	97
Percent Similarity:	53.74%	Conservative:	25
Best Local Similarity:	42.73%	Mismatches:	89
Query Match:	13.21%	Indels:	16
DB:	4	Gaps:	5

US-10-023-523-44 (1-546) x AAH07907 (1-710)

QY 28 GluGlyAlaGlnGluArgProSerGlnAlaAlaProAlaValGluAlaGluGlyProGly 47

Db 70 GAGCGAGCGGGAGAGCGCGCGCGCGCGAGAGCGAGCGAGCGCGCGCGAGCGCGG 129

QY 48 SerSerGlnAlaProArgGly-----ProGluGlyAlaGlnAlaArgThrAlaGlnSer 65

Db 130 CGCGCAGCGAGCGCGCGCGAGAGTTTGAATTCGACAAATGGAAGAAGCTGGAATTTGT 189

QY 66 GlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGluAlaGluSerThr 85

Db 190 GGGCTAGGGGTGAAGCAGCATATGTTGTGTAATCTCAATCAATCAATGATATCTTCAACAT 249

QY 86 TyrCysValAspAsnGlnGlyProGlyGluAspGlyAlaGlnGlyProAla 105

Db 250 CAA-----GGCTCAATTTGTGTCGACAGTAACAGCATTCATTGGAAGAGATCAA 303

QY 106 GluProGluAspAlaGluSerArgThrTyrValAlaArgAsnGlyGluProGluPro 125

Db 304 GCGAGTGACTTTATACAGAGAACAGGAATTTGGTGAGC-----CCA 345

QY 126 ThrProValValAsnGlyGluGluGluProSerLysGlyAspProAsnThrGluGlu 145

Db 346 GCATCTGACCGCAAGATCAAGAGAGGAATCCCTGGGGGA-----GAGCT 393

QY 146 ArgGlnSerAspGluValGlyAspArgAspHisArgProGlnGluLysLysLysAla 165

Db 334 CCAACAGATCCCCCTGATGGTCAGCAGATTCAGAGTGCAACAGGAAACAAAGAAAAACT 453

QY 166 LysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAenThrLeuSerThrPro 185

Db 454 -----TTAGAAAGAAAGTTTATTACTGATGCAAGCCCTAAACACCCCTTCAACCCCA 507

QY 186 GluGluLysLeuAlaAlaLeuCysLysLysTyrAlaGluLeuGluGluHisArgAsn 205

Db 508 GAGGAGAGCTGGGAGCTCTCTGTAAAGAAATATCTCTCTTGGAGGAGAGCAAGAGT 567

QY 206 SerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAsp 225

Db 568 GTTCAGAGCAATTCAGATCTCTGACAGAGAGAGAGCCAGATTCGAAAGAGAAAGTT 627

QY 226 HisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCys 245

Db 628 CACTTCAGAGTGAACATAGCAGGCTATCTTGGCAAGAGCAAGCTAGATCTCTNTGC 687

QY 246 ArgGluLeuGlnArgHisAsn 252

Db 688 ANAGACTTCANCGTCACAAAT 708

RESULT 42

AAAS35582

ID AAS35582 standard; cDNA; 405 BP.

XX AC AAS35582;

XX DT 17-DEC-2001 (first entry)

XX DE Human cardiovascular system antigen cDNA polynucleotide SEQ ID No 467.

XX KW Chicken; sheep; immunosuppressive; human; mouse; rabbit; goat; horse; cat;

XX KW antirheumatic; antiproliferative; cytostatic; cardiac; neuroprotective;

XX KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;

XX KW ophthalmological; vulnary; gene therapy; autoimmune disease; neoplasm;

XX KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ss;

XX KW cerebrovascular disorder; nervous system disorder; bacterial infection;

XX KW fungal infection; viral infection; ocular disorder; endocrine disorder;

XX KW gastrointestinal disorder; renal disorder; respiratory disorder;

XX KW wound healing; skin aging; organ transplantation; tissue regeneration;

XX KW anti-infertility.

XX OS Homo sapiens.

XX PN WO200155321-A2.

XX FD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001340.

XX PR 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180628P.

XX PR 24-FEB-2000; 2000US-0184664P.

XX PR 02-MAR-2000; 2000US-0186350P.

XX PR 16-MAR-2000; 2000US-0189874P.

XX PR 17-MAR-2000; 2000US-0190076P.

XX PR 18-APR-2000; 2000US-0198123P.

XX PR 19-MAY-2000; 2000US-0205515P.

XX PR 07-JUN-2000; 2000US-0209467P.

XX PR 28-JUN-2000; 2000US-0214886P.

XX PR 30-JUN-2000; 2000US-0215135P.

XX PR 07-JUL-2000; 2000US-0216647P.

XX PR 07-JUL-2000; 2000US-0216880P.

XX PR 11-JUL-2000; 2000US-0217487P.

XX PR 14-JUL-2000; 2000US-0217496P.

XX PR 26-JUL-2000; 2000US-0218290P.

XX PR 26-JUL-2000; 2000US-0220963P.

XX PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 05-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231282P.
PR 08-SEP-2000; 2000US-0231283P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0232082P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234283P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235835P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236803P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.

PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 03-DEC-2000; 2000US-0251030P.
PR 03-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 05-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-451930/48.
XX P-PSDB; AAU22308.
DR New cardiovascular system related polynucleotides and polypeptides,
XX useful for diagnosing, treating and/or preventing disorders of the
XX cardiovascular system.
PS Claim 1; SEQ ID NO 467; 674pp; English.

XX Sequences AAS35126-AAS35740 represent cDNA molecules, which encode the
XX cardiovascular system antigens polypeptides of the invention.
XX Cardiovascular system antigens and their associated polynucleotides are
XX useful in the diagnosis, treatment and prevention of various types of
XX disorders in e.g. humans, mice, rabbits, goats, horses, cats, dogs,
XX chickens or sheep. A pathological condition can be determined by
XX detecting the presence or absence of a mutation in a cardiovascular
XX system antigen polynucleotide. The treatable disorders include autoimmune
XX diseases such as rheumatoid arthritis, hyperproliferative disorders such
XX as neoplasms of the breast or liver, cardiovascular disorders such as
XX cardiac arrest, cerebrovascular disorders such as cerebral ischaemia,
XX nervous system disorders such as Alzheimer's disease, infections caused
XX by bacteria, viruses and fungi, ocular disorders such as corneal
XX infection, endocrine disorders such as premature labour and infertility.

CC gastrointestinal disorders such as Crohn's disease, renal disorders such
 CC as glomerulonephritis and respiratory disorders such as asthma and
 CC pleurisy. The polypeptides can also be used to aid wound healing, to
 CC prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, to regenerate tissues and in chemotaxis. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences

Alignment Scores:
 Pred. No.: 2,136-12 Length: 405
 Score: 352.00 Matches: 83
 Percent Similarity: 80.00% Conservative: 5
 Best Local Similarity: 75.45% Mismatches: 16
 Query Match: 12.64% Indels: 6
 DB: 4 Gaps: 2

US-10-023-523-44 (1-546) x AAS35582 (1-405)

Qy 73 GluGluLeuSerArgGlnLeuAspIleLeuSerThrTyrcysValAspAsnGln 92
 Db 75 AAGAGACGGGCTGCCACTGGAGACATATGACCATCTGTGGCAATAACAG 134
 Qy 93 GlyGlyProGlyValAspGlyValGlnGlyGluProAlaGluProGluAspAlaGlyLys 112
 Db 135 GGGGGCCCCGGCGAGGATGGGGCCACAGGGTGCAGCGGTGAACCGGAGATGCAGAGAAG 194
 Qy 113 SerArgThrTyValAlaArgAsnGlyGluPro-GluProThrProValValAsn--G1 131
 Db 195 TCCCGGACCTATGTGGCAAGGATGGGAGCTTGAACCACTTCCAKTAGTTCAATGG 254
 Qy 131 yGlyLysGluProSerLysGlyAsp-ProAsnThrGluGluIle-ArgGlnSer-AspG1 150
 Db 255 AGAGAGGAGAACCTCCAGGGGGATTCCAAACACAGAGAGATCCCGCAGAGTTGACGA 314
 Qy 150 uValGlyAspArgAspHis---ArgArgProGlnGlyLysLysAlaLysGlyLeuG1 169
 Db 315 AGTTCGGAGCAAGGAGAACCATTCGAGGCGCCACAGAGAGAAAGCCAAAGGTTTGGG 374
 Qy 169 yLysGluIleThrLeuLeuMet 176
 Db 375 GAGAGATCACGTTGCTGATT 396

RESULT 43

ADE45661

ID ADE45661 standard; cDNA; 405 BP.

XX AC ADE45661;

XX 29-JAN-2004 (first entry)

DE Human cardiovascular system related polynucleotide #457.

XX Human; cardiovascular system related polypeptide; cancer;
 KW proliferative disorder; fetal abnormality; developmental abnormality;
 KW haematopoietic disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
 KW cardiovascular disorder; angiogenic disorder; kidney disorder;
 KW gastrointestinal disorder; pregnancy-related disorder;
 KW endocrine disorder; gene; ss.

OS Homo sapiens.

XX US2003059908-A1.

XX 27-MAR-2003.

XX 07-MAR-2002; 2002US-00091504.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226688P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228324P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0232403P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0234984P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235835P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.

US-10-023-523-44 (1-546) X ABO55358 (1-725)

Qy	363	LeuLysGluAlaValGluSerGlnArgVetCysGluLeuMetLysGlnGlnGluThrHis	382
Db	334	TTAAANGAAGCCACAGAATCGAGGCACAATAACGAACAAAATGAAACACAGCANGAAGTACAA	393
Qy	383	LeuLysGlnGlnLeuAlaLeuTyThrGluLysPheGluGluPheGlnAsnThrLeuSer	402
Db	394	CTRAACACAGCACGCTTCCTTTATATGGATAAGTTTGAGAAATTCCAGACTACCATTGCCA	453
Qy	403	LysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysIle	422
Db	454	AAAAGCAATGAATCGTTTACAACTTCAGACAGAAATGGAAGAAGATGACAAAGAAAAAT	513
Qy	423	LysLysLeuGluLysGluThrThrMetTyArgSerArgTrpGluSerSerAsnLysAla	442
Db	514	AAAAAACTGGAAAAAGAAACAATAATTGGCGNACCAAANGGGAACCAT-ATAAAGCA	572
Qy	443	LeuLeuGluMetAlaGluGluLysThrValArgAspLysGluLeuGluGlyLeuGlnVal	462
Db	573	CITTCGAATGSGCTGAAGCAGGAAAACAGTCCGNGATAAACAGATNCAAGGNCTTCAATA	632
Qy	463	LysIleGlnArg-LeuGluLysLeuCysArgAla-LeuGlnThrGluArgAsnAspLeuA	482
Db	633	AAACTGACCGGTTAGAGAATTGGCGANGCTTNTTCAACCGAAGGATGGGCTCA	692
Qy	482	sn 482	
Db	693	AT 694	
	RESULT 45		
	ABN74510		
ID	ABN74510 standard; cDNA; 981 BP.		
XX	ABN74510;		
DT	03-JUL-2002 (first entry)		
DE	Bovine embryonic germ (BG) cell cDNA EST #561.		
KW	Bovine; Bos taurus; EST; expressed sequence tag; totipotency;		
OS	development; gene; ss.		
FN	Bos taurus.		
PD	WO2001194550-A2.		
PF	13-DEC-2001.		
PR	07-JUN-2001; 2001WO-US018576.		
PT	07-JUN-2000; 2000US-0209874P.		
PA	06-JUN-2001; 2001US-00876143.		
FI	(INFI-) INFIGEN INC.		
DR	Eilertsen KJ, Pfister-Genskow M, Childs L;		
PS	WPI; 2002-351289/38.		
CC	An expressed sequence tag (EST), the expression of which, or its		
CC	complementary sequence, in a cell identifies the cell as a		
CC	developmentally competent or incompetent cell.		
CC	Example. 16; Page 458-459; 584pp; English.		
CC	The present invention describes an expressed sequence tag (EST), where		
CC	the EST is an isolated, enriched, or purified nucleic acid sequence		
CC	representing all or part of a gene, the expression of which, or its		
CC	complementary sequence, in a cell identifies the cell as a		
CC	developmentally competent or incompetent cell. Molecules which induce		
CC	developmental competence in a cell line are useful for inducing		
CC	totipotency in one or more cells. Molecules which induce developmental		
CC	incompetence in a cell line are useful for preventing a full term		

CC pregnancy in an animal and inhibiting totipotency. The molecules are also
CC useful for treating a disease in an animal by inducing development of one
CC or more cells of the animal into a specific cell type. The present
CC sequence represents a bovine EST which is given in the exemplification of
CC the present invention
XX
SQ Sequence 981 BP; 195 A; 212 C; 255 G; 214 T; 0 U; 105 Other;

Alignment Scores:
Pred. No.: 1,01e-09 Length: 981
Score: 313.00 Matches: 74
Percent Similarity: 62.98% Conservativeness: 9
Best Local Similarity: 56.06% Mismatches: 46
Query Match: 11.24% Indels: 4
DB: 6 Gaps: 2

US-10-023-523-44 (1-546) x ABL03681 (1-981)

QY 45 GlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAlaArgThrAlaGln 64
Db 26 GGTGGCGCGCGCTCTAGAACTAGTGGATCCGCC---GGGCTGCAGGANTT-CGGCACGAG 81
QY 65 SerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGluAspIleLeuSer 84
Db 82 GCTGNAGCCCTCCGNGATGTTCTGAGAGTTGANCNCNCNCCTGGATGACATCCTCANN 141
QY 85 ThrTyrcysValAspAsnAsnGlnGlycylProGlyGluAspGlyAlaGlnGlyGluPro 104
Db 142 ACATACCTCGTGGACCAACAGNANNNGGCGCCAGTGNAGATGTGGCANNAGGNGAGCCT 201
QY 105 AlaGluProGluAspAlaGluLysSerArgThrTyrrValAlaArgAsnGlyGluProGlu 124
Db 202 GCTGNACCCGAGATGCANAGAGCNCNGGACCTATNCCTCANGANTGGNAGCCTGAG 261
QY 125 Pro---ThrProValValAsnGlyGlyLysGluProSerLysGlyAspProAsnThrGlu 143
Db 262 CCAGATACCTCAGTAGTCNNCTGNGAGAGGAGGANNCTCCAAAGGGGGCGCGCGGTC 321
QY 144 GluIleArg-GlnSerAspGluValGlyAspArgAspHisArgProGlnGluLys 163
Db 322 TAGATCTCTGACAGAGATGAGTTTGACCGAGNACCCCAAGCCNCGGATANTAA 381
QY 163 GlyAlaLysGlyLeuGlyLysGluIleThrLeu 174
Db 382 GAAANNCTANGTGTGGGAANGGAATACCTTG 415

RESULT 46
ID ABL03681 standard; cDNA; 7078 BP.
XX AC ABL03681;
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 5525.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PA (PEXE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI, 2001-656860/75.
DR P-FSDB; ABB59578.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 5525; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 7078 BP; 1905 A; 1854 C; 2157 G; 1162 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.57e-07 Length: 7078
Score: 291.00 Matches: 127
Percent Similarity: 44.78% Conservativeness: 113
Best Local Similarity: 23.69% Mismatches: 242
Query Match: 10.45% Indels: 55
DB: 4 Gaps: 18

US-10-023-523-44 (1-546) x ABL03681 (1-7078)

QY 3 AsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerProGlyGln 22
Db 3920 AACAGCAAGAACAGCCGGAAGAGTCCAGCAAGAACAGCAGCAGCAGTCTGGCG 3979
QY 23 ProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAlaValGlu 42
Db 3980 GAGGAGAGGTGTGTCGAGAAAGAGAAATGAAGAGAGGGCTGTCCGATGGTCAAG 4039
QY 43 AlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAlaArgThr 62
Db 4040 AAAACCATCAACAGGAGCAGGATTCGGATGCCGATCTCATCATCATCATGCTGCTG 4099
QY 63 AlaGlnSerGlyAlaLeuArgAsp-----ValSerGluGluLeuSerArgGlnLeuGlu 80
Db 4100 GCCAACAGAGCAGCARTTCGGCGCGCTCATGTCGAGCAGTCCAGTCCAGGAGTAACT 4159
QY 81 AspileLeuSerThrTyrcysValAspAsnGlnGlyGlyProGlyGlu---AspGly 99
Db 4160 TTGCACGTGGAGATGAGTGGCGGCGATCGGAGCAAGGCGGAGATGAAGATCGAAGGG 4219
QY 100 AlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrrValAlaArg 119
Db 4220 CTCGATGAGGAGCCA-----CCCAACCAATCTCCAAAGACAAAGAGAGCAAAAA 4273
QY 120 AsnGlyGluProGluProThrProValValAsnGlyGlyLysGluProSerLysGlyAsp 139
Db 4274 CCAGGAGATGCGATGAGCAACCATGCATCGAACAGAGAGAGAAAGCCAG----- 4327
QY 140 ProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHisArgPro 159
Db 4328 -----GAGAGGAGCTGAAGCTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4381
QY 160 GlnGluLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeu 179
Db 4382 GAGAAAG 4441
QY 180 AsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysTyrrAlaGluLeu 199
Db 4442 GAAAAAAGCTTAAG-----GAAGAGAGATCAAGGAG-----AAGGAACGAGAGAGAG 4489

Db 5241 TTGCAGTGGAGATGATCGCGGATCGGAGCAAGCGCAAGATCAAGAGATCAAGCGG 5182
 QY 100 AlaGlnGluValProAlaGluProGluAspAlaGluSerArgThrTyrValAlaArg 119
 Db 5181 CTCGATGAGGAGCA-----CCCAAAACAATGTCCAAAGACAACAAGAAAGCAAA 5128
 QY 120 AsnGluGluProGluProThrProValValAsnGluGluGluProSerLysGlyAsp 139
 Db 5127 CCAGGAGATCGATAGCAACCATGACATCATGACAGAGAGAGAAAGCAAG----- 5074
 QY 140 ProAsnThrGluGluLeuLeuArgGlnSerAspGluValGlyAspArgAspHisArgPro 159
 Db 5073 -----GAGAGAGGAGCTGAGCTAAAGAGAGAGAAAGAGAGAGAGAGAGAGAG 5020
 QY 160 GlnGluLysLysAlaLysGlyLeuGlyLysGluLeuThrLeuLeuMetGlnThrLeu 179
 Db 5019 GAGAAAGAGAAAGCTTAACTTAGGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4960
 QY 180 AsnThrLeuSerThrProGluGluLysLeuAlaLeuLeuCysLysLysTyrAlaGluLeu 199
 Db 4959 GAAAAAATTAAAG-----GAGAGAGAGATCAAGGAG-----AAGGAGAGAGAGAG 4912
 QY 200 LeuGluGluHisArgAsnSerGlnLysGlnMet-----LysLeu 212
 Db 4911 TTGAAAGAGAAAGAGATCAAGGAGAGAGCAAGAGAGAGAGAGAGAGAGAGAGAGCTC 4852
 QY 213 LeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSer 232
 Db 4851 AAGGAGCT 4798
 QY 233 LysAlaValLeuAlaArgSerLysLeu---GluSerLeuCysArgGluLeuGlnArgHis 251
 Db 4797 AAGGAG 4738
 QY 252 AsnArg-----SerLeuLysGluGluGlyValGlnArgAlaArgGluGluGlu 268
 Db 4737 GAAAAAATTCAAGAGAGAGCTTAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4678
 QY 269 LysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspLeuGlnMet 288
 Db 4677 CTCGAAAGAGAGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4618
 QY 289 GluGln-----HisAsnGluArgAsnSerLysLeuArgGlnGluAsn 302
 Db 4617 AAGGAGAGAGCTGAAAG 4558
 QY 303 MetGluLeuAlaGluArgLeuLysLysLeuLeuGlnTyrGluLeuArgGluGluHis 322
 Db 4557 GAAGAAAGCTCAAGGAGAGAGAGAGCTCTCAAGAGAGAGAGAGAGAGAGAGAGAG 4498
 QY 323 IleAspLysValPheLysHisLysAspLeuGlnGlnLeuValAspAlaLysLeuGln 342
 Db 4497 AGAGAGAGAGAGCTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4447
 QY 343 GlnAlaGlnGluMetLeuLysGluAlaGluArgHisGlnArgGluLysAspPheLeu 362
 Db 4446 GAACGGAGAGAGAGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4393
 QY 363 LeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnThrHis 382
 Db 4392 AGGAGAGAGAGCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4333
 QY 383 LeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSer 402
 Db 4332 GTAAAGCTTAAAG-----GAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4297
 QY 403 LysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysLe 422
 Db 4296 AAGGAGAGAGAGCTTAAAGCTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4237
 QY 423 LysLysLeuGluLysGluThrThrMetTyrArgSerArgTrpGluSerSerAsnLysAla 442
 Db 4236 AAGTCCCTGGAAG 4177

QY 443 LeuLeuGluMetAlaGluGluLys-----ThrValArgAsp-----Lys 455
 Db 4176 GTCTCGAGGATATCCACCAACTGCTCGGTTCCAGACTATCCAGCTTGGCAAG 4117
 QY 456 GluLeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGln 475
 Db 4116 AAGCCCACTAAAGCGAGTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4057
 QY 476 ThrGluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGln----- 493
 Db 4056 ACTCCCCCAAGAGTGACACACACTTTCAGAGACTTTCAGAGCGGCTT-AAAGCCACT 3998
 QY 494 GlySerLeuThrAspSerGlyProGluArgArgProGluGlyProGly 509
 Db 3997 GGAGCGCTTGCCACTCTCTCTGCTCTCGAACCCCTGGAGGTTAAGGA 3950
 RESULT 48
 ID AAS22920 standard; DNA; 5943 BP.
 XX AAS22920;
 XX 24-OCT-2001 (first entry)
 DE DNA encoding novel bone marrow polypeptide #14.
 KW Bone marrow; diagnostic; therapeutic; gene therapy; antigenic;
 KW haematopoiesis; myeloid; lymph cell disorder; tissue regeneration;
 KW wound healing; nutritional supplement; immune disorder;
 KW severe combined immunodeficiency; SCID; ds.
 OS Homo sapiens.
 XX WO200157187-A2.
 XX 09-AUG-2001.
 XX 05-FEB-2001; 2001WO-US003782.
 XX 03-FEB-2000; 2000US-00496914.
 PR 20-JUN-2000; 2000US-00598075.
 PR 19-JUL-2000; 2000US-00620325.
 PR 30-NOV-2000; 2000US-0250683P.
 XX (HYSE-) HYSEQ INC.
 PI Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Zhou P, Xue AJ;
 PI Ren F, Drmanac RJ;
 XX WPI: 2001-488875/53.
 DR P-PSDB; AAU14615.
 PT Nucleic acids encoding bone marrow polypeptides, useful in diagnostic and
 gene therapy.
 PS Claim 1; Page 160-161; 392pp; English.
 CC AAS22907-AAS23099 represent nucleic acids encoding novel bone marrow
 CC polypeptides. The nucleic acids and corresponding proteins may be used in
 CC the prevention, diagnosis and treatment of diseases associated with
 CC inappropriate bone marrow polypeptide expression. For example, to treat
 CC disorders associated with decreased expression by rectifying mutations or
 CC deletions in a patient's genome that affect the activity of the
 CC polypeptides by expressing inactive proteins or to supplement the
 CC patient's own production of the polypeptide. Additionally, the nucleic
 CC acids may be used to produce the polypeptides, by inserting the nucleic
 CC acids into a host cell and culturing the cell to express the protein. The
 CC nucleic acid and its complementary sequences may also be used as DNA
 CC probes in diagnostic assays to detect and quantitate the presence of
 CC similar nucleic acid sequences in samples, and therefore which patients
 CC may be in need of restorative therapy. The proteins may also be used as
 CC antigens in the production of antibodies against bone marrow proteins and

CC in assays to identify modulators of their expression and activity. The
 CC anti-bone marrow protein antibodies and antagonists may also be used to
 CC down regulate expression and activity. The antibodies may also be used as
 CC diagnostic agents for detecting the presence of the protein in samples
 CC (e.g. by enzyme linked immunosorbent assay (ELISA)). The proteins may be
 CC used to regulate hematopoiesis activity, and consequently in the
 CC treatment of myeloid or lymph cell disorders; in tissue regeneration,
 CC such as wound healing; as a nutritional supplement; and in treatment of
 CC immune disorders such as severe combined immunodeficiency (SCID)

XX
 SQ Sequence 5943 BP; 1437 A; 1594 C; 1783 G; 1129 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,38e-07 Length: 5943
 Score: 286.50 Matches: 165
 Percent Similarity: 36.71% Conservative: 107
 Best Local Similarity: 22.27% Mismatches: 238
 Query Match: 10.29% Indels: 233
 DB: 4 Gaps: 27

US-10-023-523-44 (1-546) x AAS22920 (1-5943)

QY 12 LysGlnSerHisProLysSerSerProGly-----GlnProGluAla 25
 DB 1152 AAGAGGAATGAGAGAAATGAGAAATGACCTTACATTTGACGGAAATATGTTAGCCCCCAAGGCA 1211
 QY 26 GlyProGluGlyAlaGlnGluArgProSerGlnAlaProAlaValGluAlaGluGly 45
 DB 1212 GACCTTACAGGC-----AGTGAACCTGCCAAAGCCTCTGAAAGGAAGCA 1256
 QY 46 Pro-----GlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
 DB 1257 CCAGAGACACAGTAGATGACAGAGAGAGGGTTCCAGGAGGGAAGAGGAGCCAAAGGAG 1316
 QY 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
 DB 1317 CCAAGAGAGAGGCTTCTGCTCTGGAAGAGGCGAGTTCAGACCCGACGCAAGCAACTGGAA 1376
 QY 80 ----- 80
 DB 1377 ATTAGTGAACACATGAAGGAACACACAGCTCTCAGACTCCATAGCTTCTGACCCCAAGTCC 1436
 QY 81 -----Asp 81
 DB 1437 TTCCATGGCTGACTTCGCTTTTCGACGCGGATCTCGAGACACCTGCTGGATGTGAT 1496
 QY 82 IleLeuSerThrTyrCysValAspAsnAsnGlnGlyProGlyGluAspGly----- 99
 DB 1497 GTGCTTTCCCGAGTCTCTGGGTGG-AGCTTGTCCGACAGGCCAGCAACCACTGGGAATAGA 1555
 QY 100 AlaGlnGly---GluProAlaGluProGluAspAlaGlnLysSer----- 113
 DB 1556 AGACAGGATGACAGCCAGTCCAGGACAGATGAGTGCAGAGCAAGCAGTCCAAAGGCT 1615
 QY 114 -----ArgThrTyrValAlaArgAsnGlyGluProGlu-Pr 125
 DB 1616 GGAGGAGAGGTACCATAGTTATCTCTCCACTTCCACAGAGAGGCGGCCAGAGTCC 1675
 QY 125 oThrProValValAsnGlyLysGluProSerLysGlyAspProAsnThrGluGlu 145
 DB 1676 CCTTCGAGCTGGCCACTGAAGAGAGAGCCCGCCAGGCGCCGAGGCGAGCCCGAGTG 1735
 QY 145 eArgGlnSerAspGluValGlyAsp----- 153
 DB 1736 GAAGGAGCAGAGGAGCTTGGGAGGAGCTCTGCAGCCAGCCTCAGCCTGCAGTGTCCCT 1795
 QY 153 ----- 153
 DB 1796 CCAGAGGAGCAGGCCCCCAAGCCACCTGCTGCTGTGAGAGGGAAGGCAAGGACGATTC 1855
 QY 154 -----ArgAspHisArgArgProGlnGlnLysLysAl 165
 DB 1856 CCAGGCGGAGGCTGGGCCCTGGGAGGAGGAGGAGGAGGATCCTGAGGAGAGGTGGC 1915

QY 165 alysglyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrPr 185
 DB 1916 GGTACGCCCCACCCGCCAGCTCTCTCCAGAGTGGATCCACAGAGCTGTGGGTCCCCC 1975
 QY 185 oGluGlu-----LysLeuAlaLeuCysLysLysTyrAlaGluLeuGlu 201
 DB 1976 AGAGCAGCTCTCAGAGGCTGCACCTAAAGGCCATGAAGAGCAGTGGCCCAAGTACTCGA 2035
 QY 201 uGluHisArgAsnSerGlnLysGlnMetLysLeuGlnLysLysGlnSerGlnLeuVa 221
 DB 2036 G-----CAAGACAGGACCTCTGATCCAGCAAGAGAGATGCA 2080
 QY 221 lGlnLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAla-----ArgSe 239
 DB 2081 GCAACTGGCGGAGAGCTGTGCCAAGAGGAGGAGAGATCTCCGGTTCACACGCA 2140
 QY 239 rLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGlu 259
 DB 2141 GAAAGAGCAATCTCTCAGTCTCTTGGAGGAGCGCTGCAGAAAGCCATTGAGGAGGAGGA 2200
 QY 259 yValGlnArgAlaArgGluGluGluLys-----ArgLysGluValTh 274
 DB 2201 GGCC---CGATGAGAGAGGAGGAAAGCCAGAGCTATCTGGCTCCGAGCTCAGTCCA 2257
 QY 274 rSerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGlu----- 289
 DB 2258 GTCCAGCACACAAGCAGATGAGGACCAATCAGGCTGACCAAGAGGCTTCCCTGCGAA 2317
 QY 290 -----GlnHisAsnGluArgAsnSerLysLeuArgGlnGluAs 302
 DB 2318 ACTGAGAGACAGTGGAGTCTCAACAGAGGCTGACAGGCGCCAGCTTGGACAGAAAA 2377
 QY 302 nMetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTrpGluLeuArgGluGluHi 322
 DB 2378 TAGGCAATGTCTGAGCAGCTCAAGGAGAGATAGAGGCTTCGGAGAGAGGAGCGAGCGC 2437
 QY 322 sIleAspLysValPheLysHisLysAspLeuGlnGln-----GlnLeuValAspAl 339
 DB 2438 TGCCCTGATCTGCAAGAGAGAGGCTCTGACAGCTGAGGAGCAGCTTGGAGGGGA 2497
 QY 339 alyLeuGlnGlnAlaGlnGluMetLeuLysGlu----- 350
 DB 2498 GAGGAAAGACCTCTGGCAACGCTGGAGAGGAGCACAGCTGCTGAGCTGGAGCGCTCTG 2557
 QY 351 -----AlaGluGluArgHisGlnArgGluLysAspPheLeuLeuLysGluAlaValG 368
 DB 2558 TCCTCTCATTTGAGGCCAAGCAGCGGAGGCTGCTCCAGCCTCCAGAGAGAGATACAGGA 2617
 QY 368 uSerGlnArgMetCysGluLeuMetLysGln-----GlnGluThrHisLeuLy 384
 DB 2618 AGCTCAGCAAGAGAGGAGGCGCCAGCTCGAAGTCTTGGGCAAGTGGAGCAGAGT 2677
 QY 384 sGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeu-----Se 402
 DB 2678 TCACCAAGTCTTATCAGTGGCTGTATGACGAGCTGAGCAGAGCTCAGAGTCTCTTCGCGA 2737
 QY 402 rLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysI 422
 DB 2738 GAAGCGCAGGAATGAGAGGAGCATGAGAGAGGTTGGACAGATG----- 2786
 QY 422 eLysLysLeuGluLysGluThrMetTyrArgSerArgTrpGluSerSerAsnLysAl 442
 DB 2787 ---AAGGAGGAGCACCAGCAAGTGTATGGCTAAGGCCAGAGAGCAGTATGAA----- 2834
 QY 442 aLeuLeuGluMetAlaGluGluLysThrValArgAspLysLeuGluGluGlnVa 462
 DB 2835 -----GCTGAGGAGAGAGAGCGCGGCTGAGCTTCTGGGGCACCTGACCGG 2881
 QY 462 lLysIleGlnArgLeuGluLysLeuCys---ArgAlaLeuGlnThrGluArgAsnAspLe 481
 DB 2882 AGAGCTGGAGCGCTTCAGAGGCGCCCATGAACGAGAACTGGAGACTGTGAGAGCAGGAGCA 2941

Db 3371 CGAGCAGCTCTTGGGGGCCCGCAGATGACAGAGAGATCAAGAGAGCTGCAGGCTCGGGCGGA 3430
QY 113 -----SerArgThrTyrValAlaArgAsnGlyGluProGluProThrProValValAsn 130
Db 3431 CGAGCTGGAAGAGGAGCTGAGGAGCAGAGCGGCGCAGCGCGGCGC----- 3475
QY 131 GlyGluLysGluProSerLysGlyAsp-ProAsnThrGluGluIleArgGlnSerAspG1 150
Db 3476 -----CGTGGAGAGCAGCTGCAGAGCGCGCGCGGAGCTGGAGGA 3517
QY 150 uValGlyAspArg-----AspHisArgArgProGlnG1 161
Db 3518 GTGACGAGCGCTGGAGAGGAGCGAGCGGCATCCGCGGGCAGCGAGGGGTGCGG 3577
QY 161 uLysLysLysAlaLysGlyLysGluLysGluIleThrLeuLeuMetGlnThrLeuAsnTh 181
Db 3578 CAAGCGGAGGCGGAG-----CTGGGAGG-----CTGCGCGGAGCTGGAGAGGC 3625
QY 181 rLeuSerThrProGluLysLeuAlaAlaLeuCysLysLysTyr-----Al 197
Db 3626 GCGCTGCGGACGAGCGCCACAGTGCAGCTGCGGCACCTGCGCGCAGCAGCGGAGCGCGC 3685
QY 197 aGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysG1 217
Db 3686 GGAGCTGGGGGAGCAGGTCGACAGCTGCAGCGGTGCGCGCAGAGCTGGAGAGGAGAA 3745
QY 217 nSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAl 237
Db 3746 GAGTGAGCTGCATGAGAGTGCAGCTGCGACCTGCTGCCAACGTGGAGAGCTCTGACCCGCGC 3805
QY 237 aArgSerLysLeuGluSerLeuCysArg----- 246
Db 3806 CAGGCCAGTCGACAGAGAGCTGTCGCGACCTATGAGGATCAGCTAAGCGAGGCCAAGAT 3865
QY 247 -----GluLeuGlnArgHis----- 251
Db 3866 CAAGGTGGAGGAGCTGCAGCGCAGCTGCGGACGCAAGCAGCAGCGTGGCGGCGACTACA 3925
QY 252 -----AsnArgSerLeuLysGlu----- 257
Db 3926 GACGGAACGGGGAGCTGAGTGGCTGCTAGAGAGAGGAGTGTCTGATCAGTCAGCT 3985
QY 258 -----GluGlyValGlnArgAlaArgGluG1 266
Db 3986 GAGCGCTGGAAGGCGCTGCGCGCCCAAGCGCTGGAAGAGTTCGCGCGCAGCTAGAGGA 4045
QY 266 uGluGluLysArgLysGluValThrSerHis-----PheGlnValThrLeuAsnAspIleG1 285
Db 4046 GGAAGCAGAGGCGCAGAGTGCCCTGCGCCACCGCTGAGGCTGTCGGCGCAGCTGTGA 4105
QY 285 nLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGluLe 305
Db 4106 CTTCTGCGGGAGCAACAC---GAGGAGAGCTCAGGCCCGAGCTGAGCTGCAGCGCT 4162
QY 305 uAlaGluArgLeuLysLysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAspLy 325
Db 4163 GCTGTCCAAAGGCCAATCGCGAGTGGCGCGCCAGTGGAGAGCAAGTACGAA---GCAGATGC 4219
QY 325 sValPheLysHisLysAspLeuGlnGln-----GlnLeuValAspAlaLysLeuGlnG1 343
Db 4220 CATCCAGAGCAGGAGCTGGAGAGGCCCAAAAAGCTGGCAGCTGCGCGCTGCAGGA 4279
QY 343 nAlaGlnGlu-----MetLeuLysGluAlaGluG1 353
Db 4280 GGCAGAGGAGGCGCTGGAGGCTGCCAACGCCAAGTGTCTATCGTTGGAGAGGCCAAGCT 4339
QY 353 uArgHisGlnArgGluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCy 373
Db 4340 GCGGTACACAGAGCTCAGAGATGTAACTCTGAGCTGGAGCGGCGGAGCTCAGCAGC 4399
QY 373 sGluLeuMetLysGlnGlnThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLy 393
Db 4400 TCGTGGCTGGACAAGAGCAGCGGCACTTGGACGGGCACTGGAGGAAACGGCGCGCGCA 4459

QY 393 sPheGluGluPheGluAsnThrLeu----- 401
Db 4460 GGAGGAGGAGATGACAGCGGAGCTGGAGGCGGCACAGAGGAGTCCCGTGGCTGGCGAC 4519
QY 402 -----SerLysSerSerGluValPheThrThrPheLy 412
Db 4520 CGAGCTTTCGCGCTGGGCGCAGCGCCACGAGGAGGACCTTGAAGCCCTGGAGACGCTCAA 4579
QY 412 sGlnGluMetGluLysMetThrLysLysLysLysLeuGluLysGluThrThrMetTy 432
Db 4580 GCGGGAGAACAGAACCTGCGAGGAGGAGTACCGACCTCACAGACAGGTGAGTCTC--- 4637
QY 432 rArgSerArgTyrGluSerSerAsnLysAlaLeuLeuGluMetAlaGluLysThrVa 452
Db 4638 -----AGTGGAGAGAGCATCCAGGAACCTG-----GAGAAACCC--- 4670
QY 452 lArgAspLysGluLeuGluGlyLeuGlnValLysIleGln---ArgLeuGluLysLeuCy 471
Db 4671 -----AAGAGGCGCTGGAAGCGAGAGAGTGAATCCAGGCTGCAGCTGGAGGAGCAGA 4726
QY 471 sArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaG1 491
Db 4727 GGGGCGCTTGGAGCTGGAGGAGACCAAGACGCTGCGGATCCAGCTGGAGCTCTCCA---CG 4785
QY 491 yGlyGlnGlySerLeuThrAspSerGlyProGluArgArg-----ProGlu--- 506
Db 4786 TCAAGCAGAGTGGACCGGAGCTGGCAGAGAACGAGGAGTGGCTTAACCTGAGGC 4845
QY 507 -----GlyPr 508
Db 4846 GCAACACACAGCAGAGCTGTGGAGTCCCTGCGAGGCTCCCTGGATGCAGAGACACGCGCCC 4905
QY 508 o-----GlyAlaGlnAla-----ProSerSerProArgValTh 519
Db 4906 GCAATGAGGCGCTGCGGCTCAAGAGAGATGAGGCGTGAACCTCAACGACCTGGAGCTGC 4965
QY 519 rGluAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyPro 538
Db 4966 AGCTGGGCCATGCCACCG-----TCAGGCCACAGAGGCCCA 5002
RESULT 50
ABK35570
ID ABK35570 standard; DNA; 5574 BP.
XX
AC ABK35570;
XX
DT 08-MAY-2002 (first entry)
XX
DE Gene MYH11 differentially expressed in breast cancer tissue.
XX
KW Human; diagnosis of breast cancer; endometrial cancer; breast tumour;
KW MAI; mitotic activity index; cytostatic; gene; ds.
XX
OS Homo sapiens.
XX
PN W0200210436-A2.
XX
PD 07-FEB-2002.
XX
PF 27-JUL-2001; 2001WO-US023642.
XX
PR 28-JUL-2000; 2000US-0222093P.
XX
PA (BCHM) BRIGHAM & WOMENS HOSPITAL INC.
PA (BAAK/) BAAK J.
PI Baak J, Mutter GL;
XX
XX WPI; 2002-180084/23.
DR P-PSDB; AAU84350.
XX
PT Diagnosing breast cancer comprises determining expression of nucleic acid

PT molecules or expression products that are differentially expressed in
 PT normal and malignant tissue.

PS Claim 1; Page 98-101; 219pp; English.

XX The present invention relates to a method for diagnosing breast cancer in
 CC a subject suspected of having endometrial cancer. The method comprises
 CC determining the expression of a set of human genes or expression products
 CC in an endometrial sample suspected of being cancerous. The human genes of
 CC the invention are differentially expressed in breast tumours
 CC characterised as high or low MAI (mitotic activity index). These sets of
 CC genes can be used to discriminate between high and low MAI breast
 CC tumours. The invention also provides DNA and protein microarrays for
 CC analysing the expression of the human genes and their protein products.
 CC The methods and arrays are useful for the diagnosis and prognosis of
 CC endometrial cancer, selecting and monitoring treatment regimes, and
 CC identification of compounds useful for the treatment of endometrial
 CC cancer. ABK3551-ABK3558 represent the human genes of the invention that
 CC are differentially expressed in breast cancer tissue

SQ Sequence 5574 BP; 1610 A; 1394 C; 1752 G; 818 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3,11e-07 Length: 5574
 Score: 284.00 Matches: 150
 Percent Similarity: 37.69% Conservative: 101
 Best local Similarity: 22.52% Mismatches: 231
 Query Match: 10.20% Indels: 185
 DB: 6 Gaps: 25

US-10-023-523-44 (1-546) x ABK35570 (1-5574)

QY 9 GlyAlaAlaLysGlnSerAsnProLysSerSerProGlyGlnProGluAlaGlyProGlu 28
 DB 3618 GGAGTCTTCAAGAGAAACCCGGCAGAGCTCAACGTTCTACAGAGCTGCGCCAGCT 3677
 QY 29 GlyAlaGlnGluArgProSerGlnAlaAlaProAlaValGluAlaGlyProGlySer 48
 DB 3678 GGA-----GGAGGA 3686
 QY 49 SerGluAlaProArgLysProGluGly---AlaGlnAlaArgThrAlaGlnSerGlyAla 67
 DB 3687 GCGGAACAGCTCAAGACAGCAGCTGGCAGGAGATGGAGGCGCAAGCAAGCACTGGAGCG 3746
 QY 68 LeuArgAspValSerGluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyCys 87
 DB 3747 CCACATCTCCACTCTCAACATCCAGCTCTCCGACTCGAAGAAAGAGAGGTTCCAGAGGAGATCGAGAACCT 3806
 QY 88 ValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluPro 107
 DB 3807 CAGCACCGTGAAGCTCTGGAAGAGGGGAAGAGAGGTTCCAGAGGAGATCGAGAACCT 3866
 QY 108 -----GluAspAlaGluLys 112
 DB 3867 CACCCAGCTACGAGAGAGCGCGCTTATGATAAAGTGGAAAGAACCAAGAACAG 3926
 QY 113 SerArgThrTyValAlaArgAsn-----GlyGluProGluProThrProVal 128
 DB 3927 GCTTCAGCAGGAGCTGGACAGCCTGTTGTTGATTGGACACACCGCCCACTCGTGTC 3986
 QY 129 ValAsnGlyGluLysGluPro-----SerLysGlyAspProAsnThr 142
 DB 3987 CAACCTGGAAAGACAGCAGGAAATTTGATCATGTTGTTAGCGAGGAGAAATCATCTC 4046
 QY 143 GluGluIleArgGlnSerAspGluValGly-AspArgAspHisArgArgProGlnGluLys 162
 DB 4047 TTCGAATACGC-----GGATGAGAGGGACAGAGCTGAGCGCAGAGCCAG 4091
 QY 162 sLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThrLe 182
 DB 4092 GGAGAAG----- 4098
 QY 182 uSerThrProGluGluLysLeuAlaLeuLeuCysLysLysLysTyAlaGluLeuGluGlu 202

DB 4099 -----GAAACCAAGGCCCTGTCCCTGGCTCGGGCCCTTGAAGAGCGCTTGAAGC 4148
 QY 202 uHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValG1 222
 DB 4149 CAAGAGGAAGTTCGAGCGGACCAACAAATCTCAAGCCGAATGGAAGACCTGGTCAG 4208
 QY 222 nGluLysAspHisLeuArgGlyGlyHisSerLysAlaValLeuAlaArgSerLysLeuG1 242
 DB 4209 CTCCAAGCATCAGCTGGGCAAGAACGTCATGAGCTGGGAAGTCCCAAGCGGCCCTGGA 4268
 QY 242 uSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGluValGlnAr 262
 DB 4269 GACCAGATGAGGAGATGAAGACGACGCTGGAAGACTGGAGACGAG-----CTGCA 4322
 QY 262 GAlaArgGluGluGluLysArgLysGluValThr-----SerHisPh 277
 DB 4323 AGCCAGGAGGACCAACCTCGGCTGGAAGTCAACATGCGAGCGCTCAAGGGCCAGTT 4382
 QY 277 eGlnValThrLeuAsnAspIleGlnLeuGlnMetGlnGlnHisAsnGluArgAsnSerLy 297
 DB 4383 CGAA-----AGGGATCTCCAAGCCCGGACGACAGCAATGAGGAGAGAGGCA 4433
 QY 297 sLeuArgGlnGluAsnMetGlu----- 304
 DB 4434 ACTGCAGACAGCTTCAAGAGTATGAGCGAACTGGAAGACGAGCGGAAGCAACGTGC 4493
 QY 305 -LeuAlaGluArgLeuLysLysLeuLeuGlu-----GlnTyGluLeu----- 318
 DB 4494 CTGGCAGCTGCAGCAAGAAAGAGCTGGAAGGGGACCTGAAAGACCTGGAGCTTCAGGC 4553
 QY 319 -----ArgGluGluHisIleAspLysValPheLys----- 328
 DB 4554 CGACTTCCATCAAGGGAGGGAGAGCCATCAAGCAGCTACGAACTGCAAGCTCA 4613
 QY 329 -HisLysAspLeuGlnGlnLeuValAspAlaLys----- 340
 DB 4614 GATGAAGGACTTCAAGAGAGAGTGAAGATGCCCGTCCAGAGATGATCTTTC 4673
 QY 341 -----LeuGlnGlnAlaG1 345
 DB 4674 CACAGCCAAAGAGAAATGAGAAGCAAGAGCTTGAAGCAGACCTCATGCGAGTACA 4733
 QY 345 nGluMetLeuLysGluAlaGluGluArgHisGlnArg-----GluLysAspPheLe 362
 DB 4734 AGAGACCTCGCCCGCTGAGAGGGCTCGCAACACAGCGACCTCGAAGAGGAGAACT 4793
 QY 362 uLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnThrHi 382
 DB 4794 GGCAGAGAGCTGGCCAGTAGCTGTCTGGGAAGGAACGCACTCCAGGACGAGAAAGCGCG 4853
 QY 382 sLeuLysGlnGlnLeuAlaLeuTyThrGluLysPheGluGluPheGlnAsnThrLeuSe 402
 DB 4854 CTGGAGCGCCGATCGCCCGCTGGAGAGAGCTGGAGGAGAGGAGCGGCAACATGGA 4913
 QY 402 rLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysI1 422
 DB 4914 GGCATGAGCGAC-----CGGTCCGCAAGAGCCACACAGCAGCG 4952
 QY 422 eLysLysLeuGluLysGluThrThrMetTyArgSerArgTrpGluSerSerAsnLysAl 442
 DB 4953 CGACAGCTCAGCAACAGCTGGCCACAGAGCGAGCAGCGCCCGCAAGAAATGAGATGTC 5012
 QY 442 aLeuLeuGluMet---AlaGluGluLysThrValArgAspLys-----GluLeuGluG1 459
 DB 5013 CCGCAGCAGCTCGAGCGGACAGACAGAGCTCCGGAGCAGCTCCACGAGATGGAGGG 5072
 QY 459 Y-----LeuGlnValLysIleGlnArgLe 467
 DB 5073 GCGCGTCAAGTCCAAGTTCAGTCCACCATCGCGCGCTGGAGCGCAAGATTCACAGCT 5132
 QY 467 uGluLysLeu-----CysArgAlaLeuGlnThrGluArgAsnAspLeu----- 481

Db 5133 GGAGGAGCAGGTCGAGCAGGAGCCAGAGAGAAACAGGCGGCCACCAAGTCGCTGAAGCA 5192
Qy 482 ----AsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAsp----- 498
Db 5193 GAAAGACAAGAGCTGAAGGAATCTT-GCTGCAGGTGAGGAGCAGCGCAAGATGGCCG 5251
Qy 499 -----SerGlyProGluuArgArgProGluGlyProGlyValaGlnAlaProSe 514
Db 5252 AGCAGTACAGGAGCAGGACAGAAAGGCAATGCCAGGGTCAAGCAGCTCAAGAGGCAGC 5311
Qy 514 rSerProArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAlaSerGl 534
Db 5312 TGGAGGAGGCAGAGGAGGAGTCCAGCGCATCAACGCCAACCGCAGGAAGCTGCAGCGGG 5371
Qy 534 YGln-----ThrGlyPro-----GlnGl 540
Db 5372 AGCTGGATGAGCCACGGAGAGCAACGAGGCCATGGCCCGGAGGTGAAGGCACCTCAAGA 5431
Qy 540 uProThrSerAlaArg 545
Db 5432 GCAAGCTCAGGCGAGG 5447

Search completed: June 8, 2004, 07:04:17
Job time : 773.041 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 8, 2004, 06:37:54 ; Search time 6234.35 Seconds
(without alignments)

3795.955 Million cell updates/sec

Title: US-10-023-523-44

Perfect score: 2785

Sequence: 1 MKQDKKNGAAKQSNPKSP.....APSTEASGQTGPQPTSARA 546

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 50 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/sgn2_1/USPTO.spool/US10023523/runat_08062004_063748_29937/app.query.fasta_1.1422
-DB=genEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=50
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10023523 @CGN 1 1 7704 @runat_08062004_063748_29937 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mui.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rtd.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	2785	100.0	1638	6	AR409338 Sequence
2	2785	100.0	1638	6	AX239603 Sequence
3	2785	100.0	1893	9	AF516206 Homo sapi
4	2780	99.8	4851	9	AL832338 Homo sapi
5	2777	99.7	4880	9	AL832636 Homo sapi
6	2741	98.4	5023	9	AL832637 Homo sapi
7	2693	96.7	4697	6	AR199536 Sequence
8	2693	96.7	4697	6	AR374688 Sequence
9	2693	96.7	4697	6	AR409323 Sequence
10	2693	96.7	4697	6	AX239574 Sequence
11	2693	96.7	4697	6	BD056450 Novel low
12	2552.5	91.7	4722	6	AR199533 Sequence
13	2552.5	91.7	4722	6	AR374685 Sequence
14	2552.5	91.7	4722	6	AR409320 Sequence
15	2552.5	91.7	4722	6	AX239571 Sequence
16	2552.5	91.7	4722	6	BD056447 Novel low
17	2456	88.2	4702	10	BC060227 Mus muscu
18	1727	62.0	1793	6	AX578043 Sequence
19	1688	60.6	1793	9	HL5444 Human high
20	1520	54.6	181547	2	AC108815 Mus muscu
21	1286	46.2	2045	6	AX878120 Sequence
22	1286	46.2	2045	6	BD157034 Primer fo
23	1286	46.2	2045	9	AK002071 Homo sapi
24	1276	45.8	4625	9	BM647633 Homo sapi
25	1276	45.8	4640	9	AL831988 Homo sapi
26	1275	45.8	2869	9	HSM803295
27	1275	45.8	4631	9	AB085905
28	1270	45.6	242048	2	HSM805276
29	1268	45.5	4647	9	HSM807623
30	1251	44.9	4704	5	D89999
31	1251	44.9	4723	6	E26533
32	1229.5	44.1	2435	10	AF422244
33	1229.5	44.1	4394	10	AF422245
34	1057	38.0	1226	10	BC059099
35	1048.5	37.6	1191	6	AR072731
36	1048.5	37.6	1191	6	BD006992
37	1008	36.2	4154	9	HSM803629
38	925.5	33.2	823	6	AX033282
39	903.5	32.4	2194	6	AX834462
40	903.5	32.4	2194	9	AK096939
41	805	28.9	529	6	BD060509
42	771	27.7	141790	10	AL671759
43	747.5	26.8	115756	9	HSDJ62215
44	740.5	26.6	22255	6	AR409342
45	740.5	26.6	22255	6	AX239608
46	707	25.4	3554	9	BC048565
47	689.5	24.8	2181	3	AK117024
48	666	23.9	3495	9	BC029686
49	597.5	21.5	1841	3	AY047502
50	568	20.4	245095	2	AC111774


```

QY 481 LeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSerGly 500
Db 1441 CTGACACAGAGGTACAGACCTGAGTGTGTGTGGCCAGGGCTCCTCACTGACAGTGGC 1500
QY 501 ProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGlu 520
Db 1501 CCTGAGAGGAGGCCAGAGGGCTCGGGCTCAAGCACCAGCTCCCGGAGGTTCACAGAA 1560
QY 521 AlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGlu 540
Db 1561 GGCCTTGTCTACCCAGGACACCGAGCACAGAGGATCAGGCCAGACTGGGGCTCAGAG 1620
QY 541 ProThrSerAlaArgAla 546
Db 1621 CCCACCTCCGCCAGGGCC 1638

RESULT 3
AF516206 1993 bp mRNA linear PRI 15-FEB-2003
LOCUS Homo sapiens taxilin mRNA, complete cds.
DEFINITION AF516206
ACCESSION AF516206
VERSION AF516206.1 GI:28394223
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1993)
AUTHORS Nogami,S., Satoh,S., Nakano,M., Shimizu,H., Fukushima,H.,
Matsuyama,A., Terano,A. and Shirataki,H.
TITLE Taxilin; a novel synaxin-binding protein that is involved in
Ca2+-dependent exocytosis in neuroendocrine cells
Genes Cells 8 (1), 17-28 (2003)
JOURNAL 22448667
PUBMED 12558796
REFERENCE 2 (bases 1 to 1993)
AUTHORS Shirataki,H., Nogami,S. and Satoh,S.
DIRECT SUBMISSION
JOURNAL Submitted (29-MAY-2002) Division of Molecular and Cell Biology,
Institute for Medical Sciences, Dokkyo University School of
Medicine, 880 Kitakobayashi, Mibu-machi, Tochigi 321-0293, Japan
Medicine, Location/Qualifiers
FEATURES
source
1..1993
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosomes="1"
/map="lp36.ll-p34.2"
/tissue_type="brain"
79..179
/codon_start=1
/product="taxilin"
/protein_id="AA042465.1"
/db_xref="GI:28394224"
/translation="MKNDQKNGAAKQKNPKSSPQQPEAGPEGAQERPSQAAPAAVEAE
GFSSQAPKPGQAQARTQAQALRDVSELSRQLEDILSTYCVNNQGGPCEGAQG
EAPEDAKSEITVYARNGEPTPVNGEKEPSKGDENTBEIROSDEVDGRHRPQ
EKKAAGLKELTLMQTLNTLSTPEELKALCKYAELEHRSOKOMKLOKQOS
QLVQKDHRLGHSKAVLARSKLSELCELOPHNSLKEEGVQPARREEKEKVTSH
FOVTLNDIQLEQHNHNSKLQENMELAEKLLKLEQYELREBHIDKVPKXDLQ
QLVDAKLOQAQAKAEERHOREKDFLLKEAVESQRMCELMKQOETHLKOOLALYE
KFEFQNTLGSSESVFTTQKEMXNTKKIKLEKETTYSRWSNKKALLEWAEK
TVRDKELGLQVRIORLEKLCROTERNLNRVQDLSAGQGSGLTDSGPERPEGP
GAQAPSSPRVTEAPCPVGPAPSTASQITGPQPTSA"
ORIGIN
Alignment Scores:
Pred. No.: 2,35e-123 Length: 1993
Score: 2785.00 Matches: 546
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0

```

```

Db: 9 Gaps: 0
US-10-023-523-44 (1-546) x AF516206 (1-1993)
QY 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerSerPro 20
Db 79 ATGAGAGACCAACAGACCAAAAGAACGGGGCTGCCAAACATCCATCCAAAGACAGCCCA 138
QY 21 GlyLysProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
Db 139 GGACCAACCGGAGACAGACCCGAGGAGCCAGAGCGGCCAGCCAGCGCGCTCCTGCA 198
QY 41 ValLysAlaLysGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
Db 199 GTAGAAGCAGAGGTCCCGGACAGCCAGGCTCTCGGAAGCCGAGGGGGCTCAGGC 258
QY 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
Db 259 AGAACGGCTCAGTCTGGGGCCCTTCGTGATGTCTCTGAGAGCTGAGCGCCCAACTGGAA 318
QY 81 AspLysLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAla 100
Db 319 GACATATCGACACATACTGTGTGGCAATAACACGGGGGGGGCCCCGGCAGGATGGGCA 378
QY 101 GlnGlyLysProAlaGluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsn 120
Db 379 CAGGTGAGCGGGCTGAACCCGAGATGCAGAGAGTCCCGGACCTATGTGGCAAGGAAT 438
QY 121 GlyLysProGluProThrProValValAsnGlyGlyLysGluProSerLysGlyAspPro 140
Db 439 GGGGAGCCTGAACCAATCCAGTAGTCAATGGAGAGAAGAACCCCTCCAAAGGGGGATCCA 498
QY 141 AsnThrLysLysLysLysLysGlyLeuValGlyAspArgAspHisArgArgProGln 160
Db 499 AACACAGAGAGATCCGGCAGATGACAGGTCCGAGACCGAGACCATCGAAGGCCACAG 558
QY 161 GluLysLysLysLysLysGlyLeuGlyLysGluLeuThrLeuLeuMetGlnThrLeuAsn 180
Db 559 GAGAAGAAAAAGCCAAAGGGTTGGGGAAGGAGATCAGCTTGTGTGATGCAGACATTTGAAT 618
QY 181 ThrLeuSerThrProGluGlyLysLeuAlaAlaLeuCysLysLysTyrAlaGluLeuLeu 200
Db 619 ACTCTGAGTACCCAGAGGAGAGCTGGCTCTGTGCAAGAAGATGCTGAACTGCTG 678
QY 201 GluLysHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeu 220
Db 679 GAGGAGCACCGGAATTCACAGAGAGCAGATGAGCTCTACAGAAAAAGCAGAGCCAGCTG 738
QY 221 ValGlnLysLysLysLysLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLys 240
Db 739 GTGCAAGAGAGAGCACCACCTGCGCGGTGAGCAGCAGAACGCCCTCTGCGCCGAGCAAG 798
QY 241 LeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGlyVal 260
Db 799 CTTGAGAGCCTATGCCGTGAGCTGCAGCGGCACACCGCTCCCTCAAGGAAGAGGTGTG 858
QY 261 GluArgAlaArgGluGluGluLysArgLysGluValThrSerHisSphGlnValThr 280
Db 859 CAGCGGGCCCGGAGGAGGAGAGCGCAGAGGTGACCTCGCACCTTCACAGGTGACA 918
QY 281 LeuAsnAspLysGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGln 300
Db 919 CTGATGACATTAGCTGCAGATGGAAACAGCACAAATGAGCGCACTCCCAAGCTGCGCAA 978
QY 301 GluAsnMetGluLeuAlaGluArgLeuLysLysLeuLysGlnGlnTyrGluLeuArgGlu 320
Db 979 GAGAACATCGAGCTGCTGAGAGGCTCAAGAGCTGATTGAGCAGTATGAGCTGCGCGAG 1038
QY 321 GluHisLysLysLysValPheLysHisLysAspLeuGlnGlnLeuValAspAlaLys 340
Db 1039 GAGCATATCCACAAGGCTTCAACACAGAGACCTACACAGCAGGTGTGGATGCCAAG 1098
QY 341 LeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGlnArgGluLysAsp 360

```



```

Db      523 AACACAGAGATCCGACAGATGACGAGTTCGAGACCGAGACCATCAAGGCCACAG 582
Qy      161 GluLysLysLysAlaLysGlyLeuGlyLysGluLysThrLeuLeuMetGlnThrLeuAsn 180
Db      583 GAGAGAAAAAGCCAAAGGGTTTGGGGAAGGAGATCACGTTGCTGATGACAGCATTTGAAT 642
Qy      181 ThrLeuSerThrProGluGluLysLeuAlaLeuLysLysLysLysLysLysLysLysLys 200
Db      643 ACTCTGATACCCAGAGGAGAGAGCTGGCTGCTCTCTGTCGAGAGAGTATGCTGACTGCTG 702
Qy      201 GluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeu 220
Db      703 GAGGAGCACCGGAATTCACAGAGCAGATGAAGCTCTCTACAGAAAAAGCAGAGCCAGCTG 762
Qy      221 ValGluLysLysAspHisLeuArgGlyGlyHisSerLysAlaValLeuAlaArgSerLys 240
Db      763 GTCAAGAGAGAGACCACTCGCGGTGAGCAGCAAGAGCGCTCTCTGCGCCGACAG 822
Qy      241 LeuGluSerLeuLysCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGlyVal 260
Db      823 CTTCAGAGCTATGCGCTGAGCTGACGCGCACAACCGCTCTCTCAAGGAGAGGTGTG 882
Qy      261 GlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnValThr 280
Db      883 CAGCGGCGCGGAGGAGGAGGAGCGCAAGAGGTGACCTCGCCTCCAGCTTCCAGGTGACA 942
Qy      281 LeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGln 300
Db      943 CTGAATGACATTCAGCTGCAGATGCAACAGCACAATGAGCGCACTCCAGCTGGCCCA 1002
Qy      301 GluAsnMetGluLeuAlaGluArgLysLysLysLeuLysGluGlnTyrGluLeuArgGlu 320
Db      1003 GAGAACATGAGCTGCTGCTGAGAGCTCAAGAAAGCTGATTCAGCAGTATGAGCTGCGCAG 1062
Qy      321 GluHisLysLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLys 340
Db      1063 GACATATCGCAAAAGCTTCAACACAGAGCCTACACAGCAGCTGTGTGATGCCAAG 1122
Qy      341 LeuGlnGlnAlaGlnLeuMetLysLysGluAlaGluArgHisGlnArgGluLysAsp 360
Db      1123 CTCACAGCGCCAGCAGATGCTTAAGGAGGAGGAGGAGCGCCACCGCGGAGAGGAT 1182
Qy      361 PheLeuLysLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlu 380
Db      1183 TTTCCTCTGAAGAGCAGTATGAGTCCCAAGAGATGTGTGAGCTGATGAAGCAGCAAG 1242
Qy      381 ThrHisLysLysGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThr 400
Db      1243 ACCACCTGAAGCAACAGCTTGCCTATACACAGAGAGTGTGAGAGTTCACAGAACACA 1302
Qy      401 LeuSerLysSerSerGluValPheThrPheLysGlnGluMetLysLysMetThrLys 420
Db      1303 CTTTCCAAAGCAGCGAGGTATTCCACCATTCACAGCAGAGATGGAAGATGACTAAG 1362
Qy      421 LysIleLysLysLeuGluLysGluThrMetTyrArgSerArgTyrGlnSerSerAsn 440
Db      1363 AGATCAAGAGCTGAGAAAGAAACCAACCATGTACCGTCCCGTGGGAGAGCGCAAC 1422
Qy      441 LysAlaLeuLeuGluMetAlaGluLysLysThrValArgAspLysGluLeuGlyLeu 460
Db      1423 AAGGCCCTGTTGAGATGCTGAGGAGAAACAGTCCGGGATAAAGAACTGGAGGCGCTG 1482
Qy      461 GlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAsp 480
Db      1483 CAGTAAATATCCACCGTGGAGAGAGCTGTCCGGGCACTCCACACAGAGCGCAATGAC 1542
Qy      481 LeuAsnLysArgValGlnAspLeuSerAlaGlyGlnGlySerLeuThrAspSerGly 500
Db      1543 CTGAACAGAGGGTACAGACCTGAGTGTGCTGGCCAGGGCTCCCTCACTCAGCAGTGGC 1602
Qy      501 ProGluArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGlu 520

```

```

Db      1603 CCTCAGAGGAGGCGCAGAGGGGCTTGGGGCTCAAGCACCAGCTCCCGGTCACAGAA 1662
Qy      521 AlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGlu 540
Db      1663 CGCGCTTGCTACCCAGAGCAGCGAGCGAGAGCATCAGGCCAGACTGGCGCTCAAGAG 1722
Qy      541 ProThrSerAlaArgAla 546
Db      1723 CCCACCTCCGCGAGGGCC 1740

RESULT 5
HSM803946
LOCUS
DEFINITION Homo sapiens mRNA; cDNA DKFZp451I0918 (from clone DKFZp451I0918);
complete cds.
ACCESSION AL832636
VERSION AL832636.2 GI:30266324
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4880)
AUTHORS Ansong, W., Krieger, S., Regiert, T., Rittmueller, C., Schwager, B.,
Mewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and
Wiemann, S.
TITLE Direct Submission
JOURNAL Submitted (17-JUN-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuerberg, GERMANY
COMMENT On Apr 30, 2003 this sequence version replaced gi:21733212.
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp451I0918) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heuberweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
FEATURES
source
1..4880
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp451I0918"
/db_xref="taxon:9606"
/map="1p34.2-36.11"
/clone="DKFZp451I0918"
/tissue_type="human spinal cord"
/clone_lib="451 (synonym: hlcc1). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
1..4880
/gene="DKFZp451I0918"
125..1765
/gene="DKFZp451I0918"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAD89951.1"
/db_xref="GI:30266325"
/translation="MKNQDKKGAQAQQSNPKSPGQPEAGPEGAGQRPSSQAAPAVEAE
GPSSQAPARKPEGAQTAQSGALRDVSELSRLQLEDILSTYCDNVNQGSGEDGAQG
EPAPEDAKRSRYVARNGEPPTPVNGKEPKSGDNPTEIRDSDEYGDHRRPQ
EKKAGLGKEITLLMOTLNTLSTPEKLAALCKYAELEBRNSOKMKLQKQS
OLVQVXDHARGHSAVLARSKLESICRELQRNLSLKEEGVQVAPEREKEVTS
FQVLNDIQLQMECHNERNSKLRQENNELARLKLIEQLREEDHDKVFKHLQ
QLVDAKLQAQERHOREKDFLKEAVESQRMCELMKQQLTHLQKQALITE
KFEEFQNTLSKSEVFTFEQEMKMTKKIKLEKTTMYRWRSSNKLLEMAEBK
TVADLEGLQVKIQRLKLCALQTERLKNRQVQLSAGGQGLTDSGPPRRPEGP
GAQPSRPVTEAPCYGAPSTASQTPQEPPTSARA"
polyA_signal
4833..4838
/gene="DKFZp451I0918"
polyA_site
4852

```

/gene="DKFZp451I0918"

ORIGIN

Alignment Scores:

Pred. No.: 1,426-122 Length: 4880
 Score: 2777.00 Matches: 544
 Percent Similarity: 100.00% Conservatives: 2
 Best Local Similarity: 99.63% Mismatches: 0
 Query Match: 99.71% Indels: 0
 DB: 9 Gaps: 0

US-10-023-523-44 (1-546) x HSM803946 (1-4880)

QY 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerSerPro 20
 DB 125 ATGAAGAACCAAGACAAAGAAAGACGGGGTGGCCCAACAATCCAAATCCAAAAGAGAGCCCA 184
 QY 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
 DB 185 GGACAAACCGAAGACAGGACCCGAGGGAGCCCAAGAGCGGCCAGCCAGCGGCTCTGCA 244
 QY 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
 DB 245 GTAGAAGCAGAAAGGTCCCGGACAGCCAGGCTCCTCGGAAGCCGAGGGGGCTCAAGCC 304
 QY 61 ArgThrAlaGlnSerClyAlaAlaLeuArgAspValSerGluLeuSerArgGlnLeuGlu 80
 DB 305 AGAAGCGCTCAGTCTGGGGCCCTTCGTGATGTCCTGAGGAGCTGAGCGGCCCACTGGAA 364
 QY 81 AspileLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGlyGlyAspGlyAla 100
 DB 365 GACATACTGAGCACATACCTGTGGACATAACACAGGGGGGGCCCGGCGAGGATGGGCA 424
 QY 101 GlnGlyGluProAlaGluProGluAspAlaGlyLysSerArgThrTyrValAlaArgAsn 120
 DB 425 CAGGGTGAGCCGGCTGAACCCGAGATGCGAGAAAGTCCCGGACCTATGTGCAAGGAAT 484
 QY 121 GlyGluProGluProThrProValValAsnGlyGlyLysGluProSerLysGlyAspPro 140
 DB 485 GGGGAGCTTGACCACTCCAGTAGTCATGAGAGAGAGAACCTCCAAAGGGGATCCA 544
 QY 141 AsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAspHisArgArgProGln 160
 DB 545 AACACAGAAGAGATCCGGCAGAGTACGAGGTGCGAGACCCGAGACCATCGAAGGGCCACAG 604
 QY 161 GluLysLysAlaLysGlyLysGlyLysGlyLysGlyLysLeuLeuMetGlnThrLeuAsn 180
 DB 605 GAGAGAAAAGCCCAAGGTTTGGGAAGAGAGATCACGTTCTGATGACGACATTTGAAT 664
 QY 181 ThrLeuSerThrProGluGluLysLeuAlaAlaLeuCysLysLysTyrAlaGluLeuLeu 200
 DB 665 ACTCTGAGTACCCACAGAGAGAGCTGGCTGCTCTGTCGAAGAAGTATGCTGAATGCTG 724
 QY 201 GluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeu 220
 DB 725 GAGGAGACCGGAAATTCACAGAAGCAGATGAAGCTCCTACAGAAAAGCAGAGCCAGCTG 784
 QY 221 ValGlnGluLysAspHisLeuArgGlyGlyHisSerLysAlaValLeuAlaArgSerLys 240
 DB 785 GTCCAGAGAGAGACCACTGCGCGTGAGCAGCAGCAGCGCTCTGCGCCCGCAGCAAG 844
 QY 241 LeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGluGlyVal 260
 DB 845 CTTGAGAGCCTATGCGGTGAGCTGCGAGCGGCACACCGCTCCCTCAAGAAAGAGGTGTG 904
 QY 261 GlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnValThr 280
 DB 905 CACGGGCGCGGAG 964
 QY 281 LeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGln 300
 DB 965 CTGAATGACATTCAGTGTGAGATGGAACAGCACAATGAGCGCAACTCCAAAGTGGCCCAA 1024

QY 301 GluAsnMetGluLeuAlaGluArgLeuLysLysLeuLysGlnTyrGluLeuArgGlu 320
 DB 1025 GAGAACATGAGCTGGCTGAGAGGCTCAAGAACTGATTGAGCAGTATGAGTGGCGGAG 1084
 QY 321 GluHisIleAspLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLys 340
 DB 1085 GAGCATATCGACAAAGTCTTCAAAACAAGAGGACCTACAAACAGCAGCTGGTGGATGCCAAG 1144
 QY 341 LeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluGluArgHisGlnArgGluLysAsp 360
 DB 1145 CTCAGCAGGCCCGAGGAGATGCTAAAGAGGCGAAGAGCGGACCCAGCGGGAGAGGAT 1204
 QY 361 PheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlu 380
 DB 1205 TTCTCTCTGAAAGAGCAGTAGAGTCCAGAGAGATGTGTGAGCTGATGAAGCAGCAAG 1264
 QY 381 ThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThr 400
 DB 1265 ACCCACCTGAAGCAACAGCTTGCCCTATACACAGAGAAGTTTGAGGAGTTCCAGAAACACA 1324
 QY 401 LeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLys 420
 DB 1325 CTTTCCAAAAGCAGCGAGGTATTCACCACTTCAGCAGCAGAGATGGAAAAGATGACTAAG 1384
 QY 421 LysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArgTrpGluSerSerAsn 440
 DB 1385 AAGATCAAGAAGCTGGAGAAAGAAACCCACCATGTACCGGTCCCGGTGGGAGAGCAGCAAC 1444
 QY 441 LysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGluLeuGluLysLeu 460
 DB 1445 AAGCCCTGTGTGAGTGGCTGAGGAGAAACAGTCCGGGATAAGAACTGGAGGGGCTG 1504
 QY 461 GlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAsp 480
 DB 1505 CAGTAAATCCAAACGGCTGGAGAGCTGTGCGGGCACTGCAGACAGAGCGCAATGAC 1564
 QY 481 LeuAsnLysArgValGlnAspLeuSerAlaGlyGlnGlnGlnSerLeuThrAspSerGly 500
 DB 1565 CTGAACAGAGGGTACAGGACCTAGTGTGGTGGCCAGGGCTCCCTCCTCAGTACAGTGGC 1624
 QY 501 ProGluArgArgProGluGlyProGlyValAlaGlnAlaProSerSerProArgValThrGlu 520
 DB 1625 CCTGAGAGAGGCCAGAGGGGCTGGGGCTCAAGACCCAGCTCCCGAGGTCACAGAA 1684
 QY 521 AlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGlu 540
 DB 1685 GCGCTTGTCTACCCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1744
 QY 541 ProThrSerAlaArgAla 546
 DB 1745 CCCACCTCCCGCCAGGGCC 1762

RESULT 6

HSM803947 5023 bp mRNA linear PRI 17-JUN-2003
 LOCUS Homo sapiens mRNA; cDNA DKFZp451J0118 (from clone DKFZp451J0118).
 DEFINITION AL832637
 ACCESSION AL832637.2 GI:30268326
 VERSION 1
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 5023)
 Anscorge, W., Krieger, S., Regiert, T., Rittmueller, C., Schwager, B.,
 Mewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and
 Wiemann, S.
 Direct Submission
 Submitted (17-JUN-2003) MIPS, Ingolstaedter Landstr.1, D-85764
 Neuherberg, GERMANY
 On Apr 30, 2003 this sequence version replaced gi:21733213.
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

sequenced by EMBL (European Molecular Biology Laboratories, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.
 This clone (DKFZp451J0118) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clones@rzpd.de Further information about the clone and the sequencing project is available at <http://mips.gsf.de/proj/cDNA/>.

FEATURES

Location/Qualifiers
 1..5023
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="RZPD:DKFZp451J0118"
 /db_xref="taxon:9606"
 /map="lp34.2-36.11"
 /clone="DKFZp451J0118"
 /tissue_type="human spinal cord"
 /clone_lib="451 (synonym: hlcc1). Vector pSPORT1, host DH10B, sites NotI + SalI"
 /dev_stage="adult"
 1..5023
 /gene="DKFZp451J0118"
 join(283..696,696..1922)
 /gene="DKFZp451J0118"
 /note="hypothetical protein
 The frame shift was determined manually"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="CAD89952.1"
 /db_xref="GI:30268327"
 /translation="MKNDKNGAAKQSNPKSPQPEAGPEGAQERFSQAPAAVEAE
 GPSSDAPKPEGAQRTAQSALRDVSELSRQLEDILSTYCDVNDGGGDEGAQG
 EPAEPDARKSTYVARNGEPTPVNGKEKPSKGDNTBEIROSDVGDGRDHPPO
 EKKAAGDKETLTLMTQTLTSTPEKLAALCKRYAELLEHRNSQKMLQKOS
 QLVQEDHURGHSEKAVLARKSLRESLRELRHNSLKEEGVQRAREEKKKEVISH
 FLVTNDIQLQMECHNERNLSQRLENLAERLKLIEYELRESHIDKVFHKDLQO
 QLVDAKLOAQEMKEAERHOREKDFLLKEAVASQRMCELMKQOETHLKKQALYTE
 KFEFEQNTLSKSESVFTFKQEMKMTKKIKLEKETTYRWRSESSNKLLEWAEK
 TVRDKLEGLQVKIQRLEKLCRALOTERNLNRVQDLSAGQGSLTDSGPERRPEGP
 GAQAPSSPRVTEPCYPGAPSPSTEASQITGPQPTPSARA"
 4991..4996
 /gene="DKFZp451J0118"
 5010
 /gene="DKFZp451J0118"
 ORIGIN
 Alignment Scores:
 Pred. No.: 7,38e-121 Length: 5023
 Score: 2741.00 Matches: 542
 Percent Similarity: 99.45% Conservative: 1
 Best Local Similarity: 99.27% Mismatches: 3
 Query Match: 98.42% Indels: 1
 DB: Gaps: 0
 US-10-023-523-44 (1-546) x HSM803947 (1-5023)
 QY 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerPro 20
 Db 283 ATGAAGACCAAGACAAAAGAACCGGGCTGCCAAACATCAATCCAAAAGACGCCA 342
 QY 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
 Db 343 GACAAACCGGAGCAGACCCGAGGAGCCGAGGAGCCCGCCAGCCAGCGGCTCTGCA 402
 QY 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
 Db 403 GTGAAGACGAAGAGTCCCGGACGACGACAGGCTCTCGGAAGCCGAGGGGCTCAAGCC 462
 QY 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
 Db 463 AGAACGGCTCAGTCTGGGGCCCTTGTGATGTCTCTGAGGAGCTGAGCGGCCCACTGGAA 522
 QY 81 AspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAla 100

Db 523 GACATCTGAGCACAATCTGTGTGGACAATAACAGGGGGGGCCCCCGGAGGATGGGGCA 582
 QY 101 GlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrValAlaAlaArgAsn 120
 Db 583 CAGGGTGAGCGGGCTGAACCGAGAGATGCAGAGAAGTCCCGGACCTATGTGGCAAGGAAT 642
 QY 121 GlyGluProGluProThrProValValAsnGlyGluLysGluProSerLysGlyAspPro 140
 Db 643 GGGGAGGCTGAACCAACTCCAGTAGTCAATCGAGAGAAGAACCCCTCAA-GGGGATCCA 701
 QY 141 AsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAspHisArgAspProGln 160
 Db 702 AACACAGAGAGATCCGGCAGAGTGCAGGTGGAGACCGAGACCATCGAAGCCACAG 761
 QY 161 GluLysLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsn 180
 Db 762 GAGAAGAAAAGCAAGGGTTTGGGCAAGAGATACGTTGCTGATGCAGACATTGAAT 821
 QY 181 ThrLeuSerThrProGluGluLysLeuAlaAlaLeuCysLysLysTyrAlaGluLeu 200
 Db 822 ACTCTGAGTACCCAGAGGAGAGCTGGCTGCTGTGTCAGGAAGTATGCTGAACCTGTG 881
 QY 201 GluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeu 220
 Db 882 GAGGAGCACCGGAATTCACAGAGCAGATCAAGCTCCTCAGAAAAGCAGAGCCAGCTG 941
 QY 221 ValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLys 240
 Db 942 GTGCAAGAGAGAGGACCACTGCGCGGTGAGCACAGCAAGCCGCTGCGCCCGCAGCAAG 1001
 QY 241 LeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGluGlyVal 260
 Db 1002 CTTGAGAGCCTATGCCCTGAGCTGCAGCGGCACAAACCGCTCCCTCAAGGAGAGGTGTG 1061
 QY 261 GlnArgAlaArgGluGluGluGluLysArgLysGluValThrSerHisPheGlnValThr 280
 Db 1062 CAGCGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1121
 QY 281 LeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGln 300
 Db 1122 CTGAATGACATTACGTGTCAGATGGAAACAGCACAAATGAGCGCAACTCCAAAGCTGCCCAA 1181
 QY 301 GluAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTyrGluLeuArgGlu 320
 Db 1182 GAGAACTGAGCTGGCTGAGAGGCTCAAGAGCTGATTGAGCAGTATGAGCTGCGCGAG 1241
 QY 321 GluHisIleAspLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLys 340
 Db 1242 GAGCATATCCACAAAGTCTTCAAAACACAAAGGACCTACACAGCAGCTGGTGGATGCAAG 1301
 QY 341 LeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGlnArgGluLysAsp 360
 Db 1302 CTCAGCAGCGCCCGAGAGATGCTAAAGGAGGACAGAGAGCGGCACCGCGGAGAGAGAT 1361
 QY 361 PheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlu 380
 Db 1362 TTTCTCTCGAAGAGGAGCAGTAGCGTCCAGAGGATGTGTGAGCTGATGAAGCAGCAAGAG 1421
 QY 381 ThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThr 400
 Db 1422 ACCACCTGAGGACACAGCTTGCCCTATACACAGAGAAGTTTGAGGAGTTCAGACACACA 1481
 QY 401 LeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLys 420
 Db 1482 CTTTCCAAAAGCAGCGAGGTATTTCACCAATTCAAGCAGGAGATGGAAGAGATGAGTAAG 1541
 QY 421 LysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArgTrpGluSerSerAsn 440
 Db 1542 AAGATCAAGAGCTGAGAGAAGAAACCACTGACCGTCCCGTCCGGTGGAGAGCAGCAAC 1601
 QY 441 LysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGluLeuGluGlyLeu 460

Db 1602 AAGCCCTGCTTGAGTGGCTGAGGAGAAACAGTCCGGGATGAAGAACTGGAGGGCGCTG 1661
 Qy 461 GlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAsp 480
 Db 1662 CAGTAAATCCACCGCTGGAGAGCTGTGCCGGGCACTGCAGACAGAGCGCAATGAC 1721
 Qy 481 LeuAsnLysArgValGlnAspLeuSerAlaGlyGlnGlySerLeuThrAspSerGly 500
 Db 1722 CTGAACAGAGGGTACAGAGCTGAGTGTGGTGGCCAGGGCTCCCTCACTACAGAGTGGC 1781
 Qy 501 ProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGlu 520
 Db 1782 CCTGAGAGAGCCAGAGGGGCTTGGGCTCAAGACCCAGCTCCCCAGGGTCAAGAA 1841
 Qy 521 AlaProCysTyrProGlyAlaProSerThrGluLysSerGlyGlnThrGlyProGlnGlu 540
 Db 1842 GCGCCTTCTACCCAGGAGCACCAGACAGACAGATCAGGCCAGACTGGGCGCTCAAGAG 1901
 Qy 541 ProThrSerAlaArgAla 546
 Db 1902 CCCACCTCCGCGAGGGCC 1919

RESULT 7
 LOCUS AR199536 4697 bp DNA linear PAT 20-APR-2002
 DEFINITION Sequence 17 from patent US 6355451.
 ACCESSION AR199536
 VERSION AR199536.1 GI:20249610
 KEYWORDS
 SOURCE Unknown.
 ORGANISM
 Unclassified.
 REFERENCE 1 (bases 1 to 4697)
 AUTHORS Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.
 TITLE Low density and lipoprotein binding proteins and their use in
 diagnosing and treating atherosclerosis
 JOURNAL Patent: US 6355451-A 17 12-MAR-2002;
 FEATURES
 source Location/Qualifiers
 1..4697
 /organism="unknown"
 /mol_type="unassigned DNA"

Alignment Scores:
 Pred. No.: 1,28e-118 Length: 4697
 Score: 2693.00 Matches: 529
 Percent Similarity: 99.81% Conservatives: 0
 Best Local Similarity: 99.81% Mismatches: 1
 Query Match: 96.70% Indels: 0
 DB: Gaps: 0

US-10-023-523-44 (1-546) x AR199536 (1-4697)

Qy 17 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnLysArgProSerGln 36
 Db 3 AAAGCAGCCAGACACCGAAGCAGACCGGAGGCCAGAGCGGCCAGCCAG 62
 Qy 37 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 56
 Db 63 GCGGCTCCTGCTAGAGCAGAGAGTCCGCGCAGCAGCCAGGCTCCTCGGAAGCCGGAG 122
 Qy 57 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluLysLeuSer 76
 Db 123 GGGGCTCAAGCAGACCGGCTGAGTCTGGGGGCGCTTCGTGATGCTCTGAGGAGCTGAGC 182
 Qy 77 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGly 96
 Db 183 CGCAACTGAAGACATACCTGAGACACATCTGTGACATATACCGAGGGGGCCCCCGGC 242
 Qy 97 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 116
 Db 243 GAGGATGGGCACAGGGTGGAGCGGCTGAACCCGAGATGCAGAGAAGTCCCGGACCTAT 302

Qy 117 ValAlaArgAsnGlyGluProGluProThrProValValAsnGlyGluLysGluProSer 136
 Db 303 GTGGCAGGAATGGGAGCGCTGAAACCACTCCAGTAGTCTATGAGAGAGAGGAACCTCC 362
 Qy 137 LysGlyAspProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAspHis 156
 Db 363 AAGGGGGATCCAAACACACAGAGAGATCCGCGAGAGTGCAGAGTCCGAGACCGAGACCAT 422
 Qy 157 ArgArgProGlnGluLysLysAlaLysGlyLeuGlyLysGluLysLeuThrLeuLeuMet 176
 Db 423 CGAAGGCCACAGAGAGAAAGAAAGCCAAAGGGTTTGGGGAAGAGAGATCACTGTCTGATG 482
 Qy 177 GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuLysCysLysLysTyr 196
 Db 483 CAGACATTTGATTAATCTGAGTACCCAGAGAGAGAGTGGTGTCTGTGTCAAGAAGTAT 542
 Qy 197 AlaGluLeuLeuGluLysArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLys 216
 Db 543 GCTGAACCTGCTGGAGGAGCACCGGAATTCACAGACAGATGAAGCTCTCTACAGAAAAG 602
 Qy 217 GlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeu 236
 Db 603 CAGAGCCAGCTGTGTGCAAGAGAGGACCACTCGCGGTGAGCACAGCAAGGCGCTCCTG 662
 Qy 237 AlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLys 256
 Db 663 GCCCGCAGCAAGCTTGAGAGCCTATGCCGTGAGCTGCAGCGGCACCAACCGCTCCCTCAAG 722
 Qy 257 GluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHis 276
 Db 723 GAAGAAGGTGTGAGCGCGGCCGAGGAGGAGAGAGGCAAGGAGGTGACTCGCAC 782
 Qy 277 PheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSer 296
 Db 783 TTCCAGGTGACACTGTAATGACATTCAGCTGCAGATGGAACAGACACATGAGCGCACTCC 842
 Qy 297 LysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuLysGluGlnTyr 316
 Db 843 AAGCTGCGCCAAAGAGAACATGGAGCTGGCTGAGAGGCTCAAGAGCTGATTTGACGATAT 902
 Qy 317 GluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnGlnLeu 336
 Db 903 GAGCTGCGCAGGAGCATATCGACAAAGTCTTCAACACAAAGACCTTACAACAGCAGCTG 962
 Qy 337 ValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGln 356
 Db 963 GTGGATGCCAAGCTCCAGAGCGGCCAGAGATGCTAAAGGAGGAGGAGGAGGAGGAG 1022
 Qy 357 ArgGluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMet 376
 Db 1023 CGGAGAGAGGATTTTCTCTGAAAGAGGCGAGTAGTCCCAGAGGATGTGTGAGTGTATG 1082
 Qy 377 LysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGlu 396
 Db 1083 AAGCAGCAAGAGAGACCCACCTGAAGCAACAGCTTCCCTTATACACAGAGAAGTTTGAGGAG 1142
 Qy 397 PheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGlu 416
 Db 1143 TTCAGAACACACTTTCCTCAAGAGGAGGAGGATTTCCACATTCACAGCAGGAGATGAA 1202
 Qy 417 LysMetThrLysLysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArgTyr 436
 Db 1203 AAGATGACTAAGAGATCAAGAGCTGAGAAAGAAACCAACCATGTACCGGTCCCGGTGG 1262
 Qy 437 GluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGlu 456
 Db 1263 CAGAGCAGCAACCAAGGCCCTCTTGGATGGCTGAGGAGAAACAGTCCCGGATTAAGAA 1322
 Qy 457 LeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThr 476
 Db 1323 CTGGAGGCGCTTGAGGTAAATAATCCACGGCTGGAGAGAGTGTGCCGGGCACCTGCAGACA 1382
 Qy 477 GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeu 496

```

Db      1383 GAGCGCATGACCTGACACAGAGGGTACAGGACTGAGTCTGCTGGTCCAGGAGCTCCCTC 1442
Qy      497 ThrAspSerGlyProGluArgArgProGluGlyProGluAlaGlnAlaProSerSerPro 516
Db      1443 ACTCACAGTGGCCCTGAGAGAGGCCACAGAGGGGCTCGGGCTCAAGCACCCAGCTCCCC 1502
Qy      517 ArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThr 536
Db      1503 AGGTCTACAGAAGCCCTTCTACCCAGGAGCCAGGAGCCAGACAGACATCAGGCCAGACT 1562
Qy      537 GlyProGlnGluProThrSerAlaArgAla 546
Db      1563 GGGCTCAAGAGCCACCTCCGCGCAGGGCC 1592

RESULT 8
AR374688
LOCUS      AR374688
DEFINITION Sequence 17 from patent US 6605588.
ACCESSION AR374688
VERSION    AR374688.1 GI:40077503
KEYWORDS   Unknown.
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 4697)
AUTHORS     Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.
TITLE       Low density lipoprotein binding proteins and their use in
            diagnosing and treating atherosclerosis
JOURNAL     Patent: US 6605588-A 17 12-AUG-2003;
FEATURES    Location/Qualifiers
            source          1..4697
                        /organism="unknown"
                        /mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.:      1,286-118      Length:      4697
Score:          2693.00        Matches:    529
Percent Similarity: 99.81%      Conservative: 0
Best Local Similarity: 99.81%    Mismatches: 1
Query Match:     96.70%         Indels:     0
DB:              6              Gaps:       0

US-10-023-523-44 (1-546) x AR374688 (1-4697)

Qy      17 LysSerSerProGlnProGluAlaGlyProGluGlyValaGlnGluArgProSerGln 36
Db      3 AAAAGCAGCCAGGACACCGGAGCAGGACCCGAGGAGGCCAGAGGCGGCCAGCCAG 62
Qy      37 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 56
Db      63 GCGGCTCTCGCAGTAGAAGCAGAGGCTCCCGGCAGCAGCGCTCTCTCGGAAGCCGGAG 122
Qy      57 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 76
Db      123 GGGGCTCAAGCCAGAACCGCTCAGTCTGGGGCCCTTCGTGATGCTCTGAGAGACTGAGC 182
Qy      77 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyProGly 96
Db      183 CGCCAACTGGAAGACATCTGAGCACATCTGTGTGGCAATTAACAGGGGGGCCCCGGC 242
Qy      97 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 116
Db      243 GAGGATGGGGCAGAGGTGAGCCGCTGAACCCGGAAGATGACAGAACTCCCGGACCTAT 302
Qy      117 ValAlaArgAsnGlyGluProGluProThrProValAlaAsnGlyGluLysGluProSer 136
Db      303 GTGGCAAGAAATGGGAGCCCTGAACCAACTCCAGTAGTCTATGAGAGAGGAACCACTCC 362
Qy      137 LysGlyAspProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAspHis 156
Db      363 ANGGGGGATCCAAACACAGAGAGATCCGGCAGAGTGACGAGGTGCGGAGCCGAGACCAT 422

```

```

Qy      157 ArgArgProGlnGlnLysLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMet 176
Db      423 CGAAGCCACACAGGAGAGAAAAAGCCAAAGGTTTGGGGAAGAGATCACGTTGCTGATG 482
Qy      177 GlnThrLeuAsnThrLeuSerThrProGluGluLysLysAlaAlaLeuCysLysLysTyr 196
Db      483 CAGCATTTGAATACTCTGAGTACCCAGAGAGAGAGTGGCTGCTGTGTGCAAGAAATAT 542
Qy      197 AlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLys 216
Db      543 GCTGAATCTGTGGAGGACACCGGAATTCACAGAGCAGATGAAGCTCCTACAGAAAAG 602
Qy      217 GlnSerGlnLeuValGlnGlnLysAspHisLeuArgGlyGluHisSerLysAlaValLeu 236
Db      603 CAGAGCCAGCTGGTGGCAGAGAGACCCCTGCGCGGTGAGCACAGCAAGGGCGTCTCTG 662
Qy      237 AlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnMetGluGlnHisAsnArgSerLeuLys 256
Db      663 GCCCGCAGAGCTTGAGAGCTTATGCGTGAGCTGCAGCGGCACAACCGCTCCCTCAAG 722
Qy      257 GluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHis 276
Db      723 GAAGAAGGTGTGCAGCGGGCCCGGAGAGAGAGAGGAGGCGCAAGGAGGTGACCTCGCAC 782
Qy      277 PheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSer 296
Db      783 TTCAGGTGACACTGAATGACATTCAGCTGCAGATGGAACAGACAAATGAGCGCACTCC 842
Qy      297 LysLeuArgGlnGlnAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTyr 316
Db      843 AAGCTGCGCCACAGAGAACATGGAGCTGGCTGAGAGGCTCAAGAAGCTGATTGAGCAGTAT 902
Qy      317 GluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeu 336
Db      903 GAGCTGCGGAGGAGCATATCGCAAGTCTTCAACACAGAGGACCTTACACAGCAGCTG 962
Qy      337 ValAspAlaLysLeuGlnGlnAlaGlnGluMetLysLysGluAlaGluArgHisGln 356
Db      963 GTGGATGCCAAGCTCCAGCAGGCCCCAGGAGATCTTAAGAGGCGCAGAGCGGCCACCCAG 1022
Qy      357 ArgGluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMet 376
Db      1023 CGGGAGAGGATTTCTCTCTGAAAGAGGCGAGTAGAGTCCCGAGAGGATGTGTGAGCTGATG 1082
Qy      377 LysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGlu 396
Db      1083 AAGCAGCAGAGAGCCACCTGAAGCAGCAGCTTGCCTATACACAGAGAAATTTGAGGAG 1142
Qy      397 PheGlnAsnThrLeuSerLysSerGluValPheThrThrPheLysGlnGluMetGlu 416
Db      1143 TTCAGAACACACTTTCCAAAAGCAGCAGGATTTCCACACATTCACAGCAGAGATGGAA 1202
Qy      417 LysMetThrLysLysLysLysLysLysGluThrThrMetTyrArgSerArgTTP 436
Db      1203 AAGATGACTAAGAAGATCAAGAAGCTTGGAGAAAGAAACCCATGATGACCGGTCCTGGTGG 1262
Qy      437 GluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGlu 456
Db      1263 GAGAGCAGCAACAAAGGCCCTGCTTGAGATGGCTGAGAGAAAACAGTCCGGGATAAGAA 1322
Qy      457 LeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThr 476
Db      1323 CTGAGGGGCTTGAGGTAAATTCACCGCTGGAGAGAGCTGTGCGGGCATCTGAGAGACA 1382
Qy      477 GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlnGlySerLeu 496
Db      1383 GAGCGCAATGACCTGAACAAAGAGGTACAGGACCTGAGTGTGTTGGTGGCCAGGCTCCCTC 1442
Qy      497 ThrAspSerGlyProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerPro 516
Db      1443 ACTGACAGTGGCCCTGAGAGGAGGCGCAGAGGGGCTGGGGCTCAAGCAGCCAGCTCCGCC 1502

```

Qy 517 ArgValThrGluAlaProCysTyrProGlyValaProSerThrGluAlaSerGlyGlnThr 536
Db 1503 AGGTCACAGAGCCCTTCTACCCAGGAGCCAGGACACAGAGCATCAGGCCAGACT 1562
Qy 537 GlyProGlnGluProThrSerAlaArgAla 546
Db 1563 GGGCTCAAGAGCCACCTCCGCCAGGGCC 1592

RESULT 9
AR409323
LOCUS Sequence 17 from patent US 6632923.
ACCESSION AR409323
VERSION AR409323.1 GI:40160111
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4697)
AUTHORS Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.
TITLE Low density lipoprotein binding proteins and their use in
diagnosing and treating atherosclerosis
JOURNAL Patent: US 6632923-A 1/14-OCT-2003;
FEATURES Location/Qualifiers
1..4697
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1.28e-118 Length: 4697
Score: 2693.00 Matches: 529
Percent Similarity: 99.81% Conservative: 0
Best Local Similarity: 99.81% Mismatches: 1
Query Match: 96.70% Indels: 0
DB: 6 Gaps: 0

US-10-023-523-44 (1-546) x AR409323 (1-4697)

Qy 17 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 36
Db 3 AAAGCAGCCCGAGACACACCGAGCAGGACCCGAGGAGCCAGGAGCCGAGCCAGCCAG 62
Qy 37 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 56
Db 63 CGGGCTCTCGAGTAAAGAGAGAGGTCCTCGGACAGCCAGGCTCTCGGAGCCGGAG 122
Qy 57 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 76
Db 123 GGGGCTCAAGCCAGACAGCGCTCAGTCTGGGCGCTTCTGTATGCTCTGAGGAGCTGAGC 182
Qy 77 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGly 96
Db 183 CGCCAACTGGAGAGACATCTGAGCAGATCTGTGGACATAACACAGGGGGGGCCCGCC 242
Qy 97 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 116
Db 243 GAGGATGGGGACAGGCTGAGCGGCTGAACCCGAGATGAGAGAGTCCCGGACCTAT 302
Qy 117 ValAlaArgAsnGlyGluProGluProThrProValValAsnGlyGlyGluProSer 136
Db 303 GTGGCAAGGAATGGGGAGCTGAACCACTCCAGTAGTCTATGGAGAGAAAGAACCTCC 362
Qy 137 LysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHis 156
Db 363 AAGGGGGATCCAAACACAGAGAGATCCCGCAGAGTGCAGGTCGAGACCGAGACCAT 422
Qy 157 ArgArgProGlnGluLysLysLysAlaLysGlyLeuGlyGluIleThrLeuLeuMet 176
Db 423 CGNAGGCCACGAG 482
Qy 177 GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuLysLysLysTyr 196

Db 483 CAGACATTGAATACTCTGAGTACCCAGAGAGAGAGCTGGCTGCTCTGTGCAAGAGTAT 542
Qy 197 AlaGluLeuLeuGluGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLys 216
Db 543 GCTGACTGCTGGAGAGCACCAGGAATTCACAGAGAGAGATGAAGCTCTCTACAGAAAAAG 602
Qy 217 GlnSerGlnLeuValGlnGlnLysAspHisLeuArgGlyGluHisSerLysAlaValLeu 236
Db 603 CAGAGCCAGCTGGTCAAG 662
Qy 237 AlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLys 256
Db 663 GCCCGCAGCAGCTTGGAGAGCTTATGCCGTGAGCTGCAGGGGCACACCGCTCCCTCAAG 722
Qy 257 GluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHis 276
Db 723 GAAGAGAGTGTGCAGCGGCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 782
Qy 277 PheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSer 296
Db 783 TTCCAGGTGACACTGAATGACATTCAGCTGAGATGGAAACACACACANTAGAGCGCACTCC 842
Qy 297 LysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuLeuGlnTyr 316
Db 843 AAGCTCGCCCAAGAT 902
Qy 317 GluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeu 336
Db 903 GAGCTCGCAGAGAGAGATATCGACAAAGTCTTCAACACAGAGAGAGAGAGAGAGAGAG 962
Qy 337 ValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluGluArgHisGln 356
Db 963 GTGGATGCCAAGCTCCAGAGGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1022
Qy 357 ArgGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMet 376
Db 1023 CGGAGAGAGAGATTTCTCTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1082
Qy 377 LysGlnGlnGlnThrHisLysLysGlnLeuAlaLeuTyrThrGluLysPheGluGlu 396
Db 1083 AAGCAGCAAG 1142
Qy 397 PheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGlu 416
Db 1143 TTCCAGACACACTTCCAAAG 1202
Qy 417 LysMetThrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 436
Db 1203 AAGATGACTAAG 1262
Qy 437 GluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGlu 456
Db 1263 GAGAGCAGCAACAGAGCCCTGCTTGAGTGGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1322
Qy 457 LeuGluGlyLeuGlnValLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 476
Db 1323 CTGAGGAGCCCTGAGAGTAAAGATCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1382
Qy 477 GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeu 496
Db 1383 GAGCGCAATGACCTGAACAG 1442
Qy 497 ThrAspSerGlyProGluArgArgProGluGluProGlyAlaGlnAlaProSerSerPro 516
Db 1443 ACTCAGAGTGGCCCTGAG 1502
Qy 517 ArgValThrGluAlaProCysTyrProGlyValaProSerThrGluAlaSerGlyGlnThr 536
Db 1503 AGGCTCAAG 1562
Qy 537 GlyProGlnGluProThrSerAlaArgAla 546
Db 1563 GGGCTCAAG 1592


```

RESULT 10
AX239574
LOCUS AX239574 4697 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 17 from Patent WO0164874.
ACCESSION AX239574
VERSION AX239574.1 GI:15797259
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE
AUTHORS Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.
TITLE Low density liprotein binding proteins and their use in diagnosing
and treating atherosclerosis
JOURNAL Patent: WO 0164874-A 17 07-SEP-2001;
Boston: Heart Foundation, Inc. (US)
FEATURES
source
1..4697
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
<3..1595
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC88297.1"
/db_xref="GI:15797260"
/db_xref="REMBL:CAC88297"
/translation="KSSPQPEAGPEGAQERPSOAPAVEAGPGSSOAPRPGGAO
RTAQSGALRDVSEELSRDLEILSYVDNNGGPGEDGAGPEAPEDAEKSTYVA
RNGEPEPTVVGVEKPEKSDPTEIRQSDVGDHRRPQEKXKAGLKEITILM
QTLNLSLPEEKALCKYAELEHRSQKMLKQKSLVQEKDHLRGHESKA
VLARKLESCLERQHRNSKEEGVQARBEERKEVTSFQVTLNDILOMEQHN
ENSKLQRNMELAEKLLEQYELREHIDKVPKHDLOOQLVDKLOQAQMLEKE
AERERKEDFELAEVSRQELMKQETHLQQLALYTEKFEFQNTLSKSEVF
TTFQMEKMTKKIKLSEKTTMYRSWESSNKALLEAEKTVDRKEGLGVKIOR
LEKLCRALQTERNDLNKLVODLSAGSGSLTDSGPERRPEGPQAQAPSSPRVTEAPCY
PGAPSTEASGGTGPQEPISARA"
ORIGIN
Alignment Scores:
Pred. No.: 1.28e-118 Length: 4697
Score: 2693.00 Matches: 529
Percent Similarity: 99.81% Conservative: 0
Best Local Similarity: 99.81% Mismatches: 1
Query Match: 96.70% Indels: 0
DB: 6 Gaps: 0
US-10-023-523-44 (1-546) x AX239574 (1-4697)
QY 17 LysSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluAtgProSerGln 36
DB 3 AAAGACGCCAGGACACCGGAGCAGAGACCCGAGGAGCCGAGGAGCCGAGCCAG 62
QY 37 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 56
DB 63 GCGGCTCTCGTAGTAAAGACAGAGTCCCGGACAGCAGCCAGGCTCTCGGAAGCCGGAG 122
QY 57 GlyAlaGlnAlaAtgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 76
DB 123 GGGGCTCAAGCCAGAACCGCTCAGTCTGGGGCCCTTCGTGTGTGTCTGTGAGGAGCTGAGC 182
QY 77 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGly 96
DB 183 CGCCAACCTGGAAGACATACCTGAGCACATCTGTGTGACATAATACAGGGGGGCCCCGGC 242
QY 97 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 116
DB 243 GAGATGGGGCACAGGGGTGAGCCGCTGAACCCGAAATGACAGAAAGTCCCGGACCTAT 302
QY 117 ValAlaArgAsnGlyGluProGluProThrProValValAsnGlyGluLysGluProSer 136

```

```

DB 303 GTGGCAAGGAATGGGAGCCTGAACCACTCCAGTAGTCTTATGGAGAGAGAACCCCTCC 362
QY LysGlyAspProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAspHis 156
DB 363 AAGGGGATCCAAACACAGAGAGATCCGCGAGAGTCCGAGGTCGAGAGCCGAGACCAT 422
QY ArgArgProGlnGluLysLysAlaLysGlyLeuGlyLysGluLysGluLysLeuMet 176
DB 423 CGAAGGCCACAGGAGAGAAAAGCCAGGGTTTGGGAGGAGATCAGCTTGCTGATG 482
QY GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysTyr 196
DB 483 CAGACATTGAATCTCTGAGTACCCAGAGAGAGAGCTGCTGCTGTGTCAAGAAAGTAT 542
QY AlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuGlnLysLys 216
DB 543 GCTGAATCTGTGGAGGAGCACCGGAATTCACAGAGCAGATGAAGCTCTTACAGAAAAG 602
QY GlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeu 236
DB 603 CAGAGCCAGCTGTGCAAGAGAGAGACCACTCGCGGTGAGCACAGCAAGGCGCTCTG 662
QY AlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLys 256
DB 663 GCCCGCAGCAAGCTTGAGAGCTTATCGCGTGTGAGCGGCGACACCGCTCCCTCAG 722
QY GluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHis 276
DB 723 GAAGAGGTGTGAGCGGGCCCGGAGGAGGAGAGCGCAAGCGAGGTGACCTCGCAC 782
QY PheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSer 296
DB 783 TTCAGGTGACACTGTAATGACATTCAGCTGCAGATGGAACAGCACCAATCAGCGCACTCC 842
QY LysLeuArgGlnGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuLeuGlnTyr 316
DB 843 AAGCTGCGCCACAGAGAACATCGAGCTGGCTGAGAGGCTCAGNAGCTGATTGAGCAGTAT 902
QY GluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeu 336
DB 903 GAGCTGCGGAGGAGCATATCGCAAGTCTTCAACACAGAGACCTACACAGCAGCTG 962
QY ValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluGluArgHisGln 356
DB 963 GTGGATGCCAAGCTCCAGCAGGCCCGAGGAGTGTCTTAAGAGGAGCAGAGCGGACCGAG 1022
QY ArgGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMet 376
DB 1023 CGGGAGAGAGATTTTCTCTGAAAGAGGCGAGTAGAGTCCAGAGGATGTGTGAGCTGATG 1082
QY LysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGlu 396
DB 1083 AAGCAGCAGAGAGACCCACCTGAAGCAACAGCTTGCCCTATATACAGAGAAAGTTTGGAGAG 1142
QY PheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGlu 416
DB 1143 TTCCAGACACACTTTCCAAAAGCAGCAGGATTTTCCACCATTCAGCAGGAGATGAA 1202
QY LysMetThrLysLysLysLysLysGluLysGluThrThrMetTyrArgSerArgTrp 436
DB 1203 AGATGACTAAAGAGATCAAGAGCTGGAGAAAGAAACCAACATGTACCGGTCCCGGTGG 1262
QY GluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGlu 456
DB 1263 GAGAGCAGCAACAGGCCCTGCTTCAGATGGCTGAGGAGAGAAACAGTCCGGATTAAGAA 1322
QY LeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThr 476
DB 1323 CTGGAGGGGCTTCAGGTAAAAATCCAAACGGCTGGAGAGAGCTGTGCCGGGCACATG 1382
QY GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeu 496
DB 1383 GAGCGCATGACTGAAACAGAGGGGTACAGGACCTGAGTGTGTGGTGGCAGGGCTCCCTC 1442

```


QY 497 ThrAspSerGlyProGluArgProGluGlyProGlyAlaGlnAlaProSerSerPro 516
 Db 1443 ACTGACAGTGGCCCTGAGAGAGGCGCAGAGGGGCTTGGGGCTCAAGACCCAGCTCCCCC 1502

QY 517 ArgValThrGluAlaProCysValProGlyAlaProSerThrGluAlaSerGlyGlnThr 536
 Db 1503 AGGGTCACAGAGCGCTTGTCTACCCAGGAGCACCGAGCAGAGATCAGGCCAGACT 1562

QY 537 GlyProGlnGluProThrSerAlaArgAla 546
 Db 1563 GGGCCTCAAGAGCCACCTCCGCCAGGGCC 1592

RESULT 11
 BD056450 4697 bp DNA linear PAT 27-AUG-2002
 LOCUS Novel low density lipoprotein binding proteins and their use in
 DEFINITION diagnosing and treating atherosclerosis,
 BD056450
 ACCESSION
 VERSION
 KEYWORDS JP 2001506983-A/8.
 SOURCE Aequorea victoria
 ORGANISM Aequorea victoria
 Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;
 Aequoreidae; Aequorea.
 1 (bases 1 to 4697)
 Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.
 TITLE Novel low density lipoprotein binding proteins and their use in
 diagnosing and treating atherosclerosis
 JOURNAL Patent: JP 2001506983-A 8 29-MAY-2001;
 BOSTON HEART FOUNDATION INC
 COMMENT PN JP 2001506983-A/8
 PD 29-MAY-2001
 PF 26-NOV-1997 JP 1998524870
 PR 27-NOV-1996 US 60/031930.03-JUN-1997 US 60/048547 PI
 ANN M LEES, ROBERT S LEES, SIMON W LAW, ANIBAL A. ARJONA PC
 A61K38/04.A61K39/17.A61K39/00.A61K49/00.A61K51/08, PC
 C07H21/00,
 PC C07K7/00, C07K14/705, C12N15/12, C12Q1/02, C12Q1/68, G01N33/566 CC
 Strandedness: Single;
 CC Topology: Linear;
 FH Key Location/Qualifiers.

FEATURES
 source
 1..4697
 /organism="Aequorea victoria"
 /mol_type="genomic DNA"
 /db_xref="taxon:6100"

ORIGIN
 Alignment Scores:
 Prd. No.: 1,28e-118 Length: 4697
 Score: 2693.00 Matches: 529
 Percent Similarity: 99.81% Conservative: 0
 Best Local Similarity: 99.81% Mismatches: 1
 Query Match: 96.70% Indels: 0
 DB: 6 Gaps: 0

US-10-023-523-44 (1-546) x BD056450 (1-4697)

QY 17 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 36
 Db 3 AAAGCAGCCAGGACAAACCGAGCAGGACCCAGGAGGCCAGAGGCCAGCCAGCCAG 62

QY 37 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 56
 Db 63 GCGGCTCTCGAGTAGAGCAGAGAGGTCGCGCAGCAGCAGGCTCTCGAGAGCTGAGC 122

QY 57 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 76
 Db 123 GGGGCTCAAGCAGCAGAGCGCTCAGTCTGGGGCCCTTCTGATGCTCTGAGAGCTGAGC 182

QY 77 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnGlnGlyProGly 96
 |||||

Db 183 CGCCAACTGGNAGACATACTGACACATATCTGTGTGACACATAACACAGGGGGCCCCCGG 242

QY 97 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 116
 Db 243 GAGGATGGGGCACAGGGTGAACCCCGCTGAACCCAGAGATGCAGAGAAGTCCCGACCTAT 302

QY 117 ValAlaArgAsnGlyGluProGluProThrProValValAsnGlyGluLysGluProSer 136
 Db 303 GTGGCAAGGAATGGGAGCCCTGAACCACTCCAGTAGTCTATGGAGAGAGGAACCCCTCC 362

QY 137 LysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHis 156
 Db 363 AAGGGGATCCAAACACAGAGAGATCCGCGAGAGTCCGAGGTCGAGACCGAGACCAT 422

QY 157 ArgArgProGlnGluLysLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMet 176
 Db 423 CGAAGGCCACAGGAGAGAAAGAAAGCCAGGGTGTGGGGAAGGAGATCACTGTCTGATG 482

QY 177 GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysTyr 196
 Db 483 CAGACATTGATCTCTGAGTACCCAGAGAGAGAGTGGCTGCTCTGTGCNAGAGTAT 542

QY 197 AlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLys 216
 Db 543 GCTGAACCTGTGGAGGAGCACCGGATTCACAGAGCAGATGAAGCTCTCTACAGAAAAG 602

QY 217 GlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeu 236
 Db 503 CAGAGCCAGCTGTGTGCAAGAGAGGACCCTCGCGGTGAGCACACAGGCGCTGCTG 662

QY 237 AlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLys 256
 Db 663 GCCCGCAGCAAGCTTGAGAGCTATGCGGTGAGTGCAGCGGCGCAACCGCTCCCTCAAG 722

QY 257 GluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHis 276
 Db 723 GAAGAGGTGTGACGCGGCGCGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 782

QY 277 PheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSer 296
 Db 783 TTCAGGTGACACTGATGATGACATTCAGCTCAGATGGAACAGCACATGAGCGCACTCC 842

QY 297 LysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTyr 316
 Db 843 AAGCTGCGCCAAAGAACACATGAGCTGGCTGAGAGGCTCAAGAGCTGATGATGAGCAGTAT 902

QY 317 GluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeu 336
 Db 903 GAGCTGCGCGAGGAGCATATCGACAAAGTCTTCAACACAGGACCTTACACAGCAGCTG 962

QY 337 ValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGln 356
 Db 963 GTGGATGCCAAGCTCCAGCAGGCCAGGAGATGCTAAAGGAGGAGGAGGAGGAGGAGGAG 1022

QY 357 ArgGluLysAspPheLeuLeuLysGluAlaValIleCysSerGlnArgMetCysGluLeuMet 376
 Db 1023 CGGGAAGAGATTTCTCTCTGAAAGAGGAGTATGAGTCCCGAGAGGATGTGTGAGCTGATG 1082

QY 377 LysGlnGlnGluThrHisLeuLysGlnLeuLeuAlaLeuTyrThrGluLysPheGluGlu 396
 Db 1083 AAGCAGCAAGAGAGCCACCTGAAAGCAACAGCTTGCCCTATACACAGAGAAGTTTGAGGAG 1142

QY 397 PheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGlu 416
 Db 1143 TTCAGAACACACTTTCCAAAGCAGCGAGTATTCCACCATTTCAAGCAGGAGATGGAA 1202

QY 417 LysMetThrLysLysLysLysLysLeuGluLysGluThrThrMetTyrArgSerArgTyr 436
 Db 1203 AAGATGACTAAGAGATCAGAGAGCTGGAGAAAGAAACCAACCATGTACCGTCCCGGTGG 1262

QY 437 GluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValaArgAspLysGlu 456
 Db 1263 GAGAGCAGCAACAGGCCCTGTGATGCTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1322

QY 457 LeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThr 476
 Db 1323 CTGAGGGCTCGAGGTAAATCCAAACGGCTGGAGAGCTGTGGCGGCACTGCAGACA 1382
 QY 477 GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeu 496
 Db 1383 GAGCGCAATGACCTGAACAGAGGGTACAGAGACCTGAGTGGTGGCCAGGGCTCCCTC 1442
 QY 497 ThrAspSerGlyProGluArgArgProGluGlyProGlyValaGlnAlaProSerSerPro 516
 Db 1443 ACTGACAGTGGCCCTGAGAGAGGCCAGAGGGCCCTGGGGCTCAGACCCAGCTCCGCC 1502
 QY 517 ArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThr 536
 Db 1503 AGGGTCACAGAGCGGCTTGTCTACCCAGGAGCACCAGAGCACAGAAGCATCAGGCCAGACT 1562
 QY 537 GlyProGlnGluProThrSerAlaArgAla 546
 Db 1563 GGGCCCTCAAGAGCCACCTCCCGCCAGGGCC 1592
 RESULT 12
 AR199533
 LOCUS 4722 bp DNA linear PAT 20-APR-2002
 DEFINITION Sequence 14 from patent US 6355451.
 ACCESSION AR199533
 VERSION AR199533.1 GI:20249607
 KEYWORDS Unknown.
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 4722)
 AUTHORS Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.
 TITLE Low density lipoprotein binding proteins and their use in
 diagnosing and treating atherosclerosis
 JOURNAL Patent: US 6355451-A 14 12-MAR-2002;
 FEATURES Location/Qualifiers
 source 1..4722
 /organism="unknown"
 /mol_type="unassigned DNA"
 ORIGIN
 Alignment Scores:
 Pred. No.: 5 6e-112 Length: 4722
 Score: 2552.50 Matches: 508
 Percent Similarity: 93.19% Conservative: 12
 Best Local Similarity: 91.04% Mismatches: 25
 Query Match: 91.65% Indels: 13
 DB: 6 Gaps: 3
 US-10-023-523-44 (1-546) x AR199533 (1-4722)
 QY 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerSerPro 20
 Db 61 ATGAAGAATCAAGACAAAGACGGGGCTGCCAAACAGCCCAACCCCAAGAGCACCCCG 120
 QY 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
 Db 121 GGACAGCCGAGAGAGGAGGAGGAGCCAGGGCGCGCCCGCGCGCGCGCGCGCGCC 180
 QY 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
 Db 181 CGAAGAACCCGAAGGT---GCCAGACGCCAGGCTCCCGGAGGCGCGGAGGCGGCTCAAGCC 237
 QY 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluLysSerArgGlnLeuGlu 80
 Db 238 AAAACTGCTCAGCTGGGGCGCTCTGTGATGCTCTGAGGAGCTGAGCCGCCAGTTGGAA 297
 QY 81 AspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAla 100
 Db 298 GACATACCTAGTACATACTGTGTGGCAACACACAGGGGGCGCGCGGTGAGGATGGGTC 357
 QY 101 GlnGlyGluProAlaGluProGluAlaGluLysSerArgThrTyrValAlaArgAsn 120

Db 358 CAGGGTGAGCCCTGAACCTGAAGATGCAGAGAGTCTCCGCCCTATGTGGCAGAGAAAT 417
 QY 121 GlyGluProGluPro---ThrProValValAsnGlyGluLysGluProSerLysGlyAsp 139
 Db 418 GGGGAGCCGGAGCGGGCACCCTCCAGTAGTCAATGGCGAGAGGAGACCTCCAAAGGCAGAG 477
 QY 140 ProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHisArgArgPro 159
 Db 478 CCGGGCAGCGAAGAGATCCCGACGAGCGATGAGTCCGAGACCGAGACCAACCGGAGGCCA 537
 QY 160 GlnGluLysLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeu 179
 Db 538 CAGGAAAAGAAAGAGCCAGGGCTGGGAAAGAGAGATCAGCTGCTGATGCAGACACTG 597
 QY 180 AsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysLysTyrAlaGluLeu 199
 Db 598 AACACGCTGAGCACCCCGAGGAGAGCTGGCGGCTCTGTCAAGAAAGTATGCGGAACCTG 657
 QY 200 LeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGln 219
 Db 658 CTCGAGGAGCACCGGAATCTCCAGAGCAGATGAAGCTGCTGCAGAAAGACAGAGCCAG 717
 QY 220 LeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSer 239
 Db 718 CTGCTGCAGAGAGAGGAGCCACCTGCTGGTGGCGAGCACAGCAAGGCCATCTCTGGCCCGCAGC 777
 QY 240 LysLeuGlnSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGly 259
 Db 778 AAGCTCAGAGACCTGTGCGGGAGCTGCAGCGGCACAAACCGCTCGCTCAAGAGAAAGGT 837
 QY 260 ValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnVal 279
 Db 838 GTGCAGCGAGCCCGAGAGGAGGAGAGAGCAAGCGAAGAGGTGACGTCAACATTCAGATG 897
 QY 280 ThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArg 299
 Db 898 ACGCTCAACGACATTCAGCTGCAGATGGAGCAGCACCAACAGCGCAACTCTCAAGCTCGCC 957
 QY 300 GlnGluAsnMetGluLeuAlaGluArgLeuLysLysLysLysLysGluGlnTyrGluLeuArg 319
 Db 958 CAGGAGAAACATGAGCTGGCGGCGGCTCAAGAGCTGATTGAGCAGTACGAGCTGCGA 1017
 QY 320 GluGluHisLysLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAla 339
 Db 1018 GAGAGCACATCGACAAAGTCTTCAAAACACAAAGATCTGCAGCAGCAGCTGGTGGACGCC 1077
 QY 340 LysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGlnArgGluLys 359
 Db 1078 AAGCTCCAGCAGGCGCCAGGAGATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1137
 QY 360 AspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGln 379
 Db 1138 GACTTCTCTCTGAAGGAGGCGCTGGAGTCCAGAGGATGTGCGAGCTGATGAAGCAACAG 1197
 QY 380 GluThrHisLysLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsn 399
 Db 1198 GAGACCCACCTCGAAGCAGCAGCTTGCCTATACACAGAAAGTTTGAGGATTTCCAGAAC 1257
 QY 400 ThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThr 419
 Db 1258 ACTCTTTCCAAAGCAGCGAGGTGTTCAACACATTCACCAAGAAATGGAAGATGACA 1317
 QY 420 LysLysLysLysLysLysGluLysGluThrThrMetTyrArgSerArgTyrGluSerSer 439
 Db 1318 AAGAGATCAAGAGCTGGAGAAAGAGACCAACCATGTACCGGGACAAAGAGCTGGAAGGC 1377
 QY 440 AsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGluLeuGluGly 459
 Db 1378 AACAGGCCCTGCTTGATGGCTGAGAGAGAAACACTCCGGGACAAAGAGCTGGAAGGC 1437
 QY 460 LeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsn 479

```
Db 1438 CTGCAGGTGAAATCCAGCGGCTGGAGAGCTGTGCGGGCACTGCAGACAGCGCAAT 1497
Qy 480 AspLeuSerLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSer 499
Db 1498 GACCTTGAAACAGAGGCTGCAGACCTGAGTCCGCTGGCGAGGGCCCGCTCCACAGC 1557
Qy 500 GlyProGluArgPro-----GluGlyPro 508
Db 1558 GGTCTGACGGAGCGCCAGAGCGCCACCACCTCCAGGAGCAGGGTGTGAGGGGCC 1617
Qy 509 GlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyrProGlyAlaPro 528
Db 1618 GGGGCTCAAGTACCAACCTCCAGAGGCCACAGACAGCTTCTGCTGCGCAGGTGCACCC 1677
Qy 529 SerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAlaArgAla 546
Db 1678 AGCACAGAGGCTACAGCCAGACAGAGCGGCCCGCCAGAGGCCACCTGCCACTGCC 1731

RESULT 13
AR374685
LOCUS AR374685 4722 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 14 from patent US 6605588.
ACCESSION AR374685
VERSION AR374685.1 GI:40077500
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4722)
AUTHORS Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.
TITLE Low density lipoprotein binding proteins and their use in
diagnosing and treating atherosclerosis
JOURNAL Patent: US 6605588-A 14 12-AUG-2003;
FEATURES
source
1. .4722
Location/Qualifiers
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 5,6e-112 Length: 4722
Score: 2552.50 Matches: 508
Percent Similarity: 93.19% Conservative: 12
Best Local Similarity: 91.04% Mismatches: 25
Query Match: 91.65% Indels: 13
DB: 6 Gaps: 3

US-10-023-523-44 (1-546) x AR374685 (1-4722)
Qy 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerPro 20
Db 61 ATGAAGATCAGACAAAAGAACCGGGCTGCCAACACGCCCAACCCCAAGAGCGCCG 120
Qy 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
Db 121 GGACACCGGAGACAGAGCGGAGGAGCCCGAGGGCGCGCCCGCGCGCGCGCCCGCC 180
Qy 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
Db 181 CGAAGAGCCGAAGGT---GCCAGCACCGAGGCTCCCGGAGCGCGAGGGGGCTCAAACC 237
Qy 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
Db 238 AAAACTGCTCAGCTGGGGCGCTCTGTGATGCTCTGAGAGCTGAGCGCGCAGTTGGNA 297
Qy 81 AspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAla 100
Db 298 GACATACTCAGTACATACCTGTGTGGCAACAACACCGAGGGGGCGCGGGTGGAGTGGGTC 357
Qy 101 GlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsn 120
Db 358 CAGGGTGAGCCCGCTGAACCTGAGATGCAGAGAGTCTCCGGCCTATGTGCAAGGAAT 417
```

```
Qy 121 GlyGluProGluPro---ThrProValValAsnGlyGluLysGluProSerLysGlyAsp 139
Db 418 GGGAGCGCGAGCGCGGCACCCCAAGTAGTCAATGGCGGAGAGGAGACCTCCAGGCGAG 477
Qy 140 ProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHisArgArgPro 159
Db 478 CCGGGCACGGAAGAGATCCGGACAGCGATGAGTCCGAGACCGAGACCGAGCCAGGCCA 537
Qy 160 GlnGluLysLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeu 179
Db 538 CAGGAAAAGAGAAAGGCAAGGGTCTGGGAAAGAGATCACGCTGCTGATGCAGACACTG 597
Qy 180 AsnThrLeuSerThrProGluGluLysLeuAlaLeuAlaLeuCysLysLysTyrAlaGluLeu 199
Db 598 AACACGCTGACACCCACAGAGGAAAGCTGGCGGCTCTGTGCAAGAGATATCGGAACTG 657
Qy 200 LeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGln 219
Db 658 CTCGAGGAGCACCGAACTCGCAGAGCAGATGAAGCTGCTGAGAAAGAGAGAGGCGAG 717
Qy 220 LeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSer 239
Db 718 CTGGTGACGAGAGGACCACTCGCTGGCGAGCAGACAGCAAGCCATCTCTGGCCCGCAGC 777
Qy 240 LysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGluGly 259
Db 778 AAGCTCGAGAGCCTGTGCGCGGAGCTGCAGCGGCACCAACCCCTCGCTCAAGGAGAAAG 837
Qy 260 ValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnVal 279
Db 838 GTGACGAGCGCCGAGAGGAGGAGAGCGCAGAGAGGTGAGCTCACACTTCCAGATG 897
Qy 280 ThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArg 299
Db 898 ACGCTCAACGACATTTCAGCTGCAGTGGAGCAGCACCAACGAGCGCAACTCCAAGCTCGC 957
Qy 300 GlnGluAsnMetGluLeuAlaGluArgLysLysLeuIleGluGlnTyrGluLeuArg 319
Db 958 CAGGAGAACATGGAGCTGGCGGAGCTCAAGAGCTGATTTGAGCAGTACGAGCTGCGCA 1017
Qy 320 GluGluHisIleAspLysValPheLysHisAspLeuGlnGlnGlnLeuValAspAla 339
Db 1018 GAAGAGCACATCGACAAAGTCTTCAAAACACAAAGGATCTGCAGCAGCAGCTGGTGGAGCC 1077
Qy 340 LysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGlnArgGluLys 359
Db 1078 AAGCTCCAGCAGCGCCCGAGGAGTCTGAAGAGCAGAGGAGCGGCGCCAGCGGGAGAG 1137
Qy 360 AspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGln 379
Db 1138 GACTTTCTCTGAGAGAGCGCTGGAGTCCAGAGGATGTCCGAGCTGATGAAGCAACAG 1197
Qy 380 GluThrHisLysLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsn 399
Db 1198 GAGACCCACCTGAAGCAGCAGCTTGCCCTATACACAGAGAAGTTTGAGGAGTTCCAGAAC 1257
Qy 400 ThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThr 419
Db 1258 ACTCTTTCCAAACACGAGGAGTGTTCACCATTCACAGGAATGGAAGATGACATGACA 1317
Qy 420 LysLysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArgTyrGluSerSer 439
Db 1318 AAGAAGATCAAGAAGCTGGAGAAAGAGACCAACATGTACCGTTCCCGGTGGGAGAGCAGC 1377
Qy 440 AsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGluLeuGluGly 459
Db 1378 AACAGGCGCTGTGTGAGTGGCTGAGGAGAAACACTCCGGGACAAAGAGCTGGAAAGGC 1437
Qy 460 LeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsn 479
Db 1438 CTGCAGGTGAAATCCAGCGCTGGAGAGAGTGTGCGCGGCACCTGCAGACAGACGCGCAAT 1497
Qy 480 AspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlnGlyGlnGlySerLeuThrAspSer 499
```

```

Db      1498 GACCTGACAGAGGTGCGAGACCTGAGTGGCGGTGGCCAGGGCCCGCTCTCCGACAGC 1557
Qy      500 GlyProGluArgPro-----GluGlyPro 508
Db      1558 GGTCTGAGCGGAGGCGACAGCCCGCCACCCCTCCAGGAGCAGGGTGTCTCAGGGGCCCC 1617
Qy      509 GlyAlaGlnAlaProSerProArgValThrGluAlaProCysTyrProGlyAlaPro 528
Db      1618 GGGGCTCAAGTACCACTCTCCAGGGCCACAGAGCCTTCTGTCGCCAGGTGCACCC 1677
Qy      529 SerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAlaArgAla 546
Db      1678 AGCACAGAGGCATCAGGCAGACAGGGCCCGCAGGACCCACCACTGCCACTGCC 1731

RESULT 14
AR409320 4722 bp DNA linear PAT 18-DEC-2003
LOCUS Sequence 14 from patent US 6632923.
DEFINITION AR409320
ACCESSION AR409320
VERSION AR409320.1 GI:40160108
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4722)
AUTHORS Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.
TITLE Low density lipoprotein binding proteins and their use in
diagnosing and treating atherosclerosis
JOURNAL Patent: US 6632923-A 14 OCT-2003;
FEATURES Location/Qualifiers
source 1..4722
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 5,6e-112 Length: 4722
Score: 2552.50 Matches: 508
Percent Similarity: 93.19% Conservative: 12
Best Local Similarity: 91.04% Mismatches: 25
Query Match: 91.65% Indels: 13
DB: 6 Gaps: 3

US-10-023-523-44 (1-546) x AR409320 (1-4722)

Qy      1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerSerPro 20
Db      61 ATCAAGAAATCAAGACAAAAGAAACCGGGCTGCCAACACAGCCCAACCCCAAAAGCAGCCG 120
Qy      21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
Db      121 GGACACCGGGAACAGACAGAGCGGAGGAGCCAGGGGCGCGCGCGCGCGCGCGCGCGCGCG 180
Qy      41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
Db      181 CGAGAGACCGAAGT---GCCAGCACCGAGGTCTCCGGGAGCGCGGAGGGGGCTCAAGCC 237
Qy      61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
Db      238 AAAACTGCTCAGCTCGGGCGCTCTGTGTGTCTCTGAGAGCTGAGCGCGCCAGTGTGGA 297
Qy      81 AspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAla 100
Db      298 GACATACTCAGTACATCTGTGTGGACACACACCGAGGGGGCGCGCGCGCGCGCGCGCG 357
Qy      101 GlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsn 120
Db      358 CAGGGTGAGCCCCCTGAACCTGAAGATGACAGAGAGTCTTCGGCGCTATGTGGCAAGGAAT 417
Qy      121 GlyGluProGluPro---ThrProValValAsnGlyGluLysGluProSerLysGlyAsp 139
Db      418 GGGGAGCCGGAGCCGGGCGCCCGGAGTGTGTCNTGGCGAGAGAGACCTCCAGGCGCAGAG 477

```

```

Qy      140 ProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHisArgGtpPro 159
Db      478 CCGGGCCAGGAGAGATCCCGACGAGCGATAGGTCCGAGACCGAGACCAACCGGAGGCCA 537
Qy      160 GlnGlyLysLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeu 179
Db      538 CAGAAAAGAAAGAGGCAAGGGTCTCGGAAAGAGAGATCAGCTGCTGATGTCAGACACTG 597
Qy      180 AsnThrLeuSerThrProGluGluLysLeuAlaAlaLeuCysLysLysTyrAlaGluLeu 199
Db      598 AACACGCTGAGACCCCGAGAGGAGAGCTGGCGCTCTGTGCAAGAGATATGGCGAACTG 657
Qy      200 LeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGln 219
Db      658 CTCGAGGAGCACCGGAACCTCGCAGAGCAGATGAAGCTGCTGCAGAAAGACGAGAGCCAG 717
Qy      220 LeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSer 239
Db      718 CTGCTGCAGAGAGAGGACCACTCGCTGGCGAGCACAGCAAGGCCATCTCTGGCCCGCAGC 777
Qy      240 LysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGly 259
Db      778 AAGCTCGAGAGCTGTGCGGGAGCTGCAGGGGCACAAACCGCTCGCTCAGAGGAAGAGT 837
Qy      260 ValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnVal 279
Db      838 GTGCAGCGAGCCCGAGAGGAGGAGAGCGCAAGAGGTGACGTCCACACTTCCAGATG 897
Qy      280 ThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGlnArgAsnSerLysLeuArg 299
Db      898 ACGTCAACGACATTCAGCTGCGAGTGCAGACAGCACAAACGAGCGCAACTCCAGAGTGGCG 957
Qy      300 GlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTyrGluLeuArg 319
Db      958 CAGGAGAACTGGAGCTGGCGGAGCGGCTCAAGAGCTGATTGAGCAGTACGAGTGCGA 1017
Qy      320 GluGluHisLysAspLysValPheLysHisLysAspLeuGlnGlnLeuValAspAla 339
Db      1018 GAAGAGCACATCCGACAAAGTCTTCAACACAAAGATCTGCAGCAGAGCTGGTGGACGCC 1077
Qy      340 LysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGlnArgGluLys 359
Db      1078 AAGCTCCAGCAGGCCCGCAGGAGATGCTGAAGAGGAGGAGCGGCACACGCGGGAGAG 1137
Qy      360 AspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGln 379
Db      1138 GACTTTCTCTGAGGAGGCGCGTGGAGTCCCGAGGATGTGCGAGCTGTGAAGCAACAG 1197
Qy      380 GluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsn 399
Db      1198 GAGACCCACCTGAGCAGCAGCTTGCCTATACACAGAGAGTTTGAGGATTCAGAAC 1257
Qy      400 ThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThr 419
Db      1258 ACTCTTCCAAAAGCAGCAGGTGTTCAACACATTCACACAGGAAATGGAAGATGACA 1317
Qy      420 LysLysLysLysLysLeuGluLysGluThrThrMetTyrArgSerArgTyrGluSerSer 439
Db      1318 AAGAAATCAAGAACTGGAGAAAGAGACCAACCATGATACCTTCCCGGTGGGAGACAGC 1377
Qy      440 AsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGluLeuGly 459
Db      1378 AACAAAGCCCTGCTTGAATGGCTGAGGAGAAAACACTCCGGGACAAAGAGCTGGAGGC 1437
Qy      460 LeuGlnValLysLysGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsn 479
Db      1438 CTGAGGTGAAATCCAGCGCGCTGGAGAGCTGTGCGGGCACTGCAGACAGAGCGCAAT 1497
Qy      480 AspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlnGlySerLeuThrAspSer 499
Db      1498 GACCTGAACAAGAGGGTGCAGGACCTGAGTGGCGGTGGCGCGCGCGCTCTCCGACAGC 1557

```

QY 500 GlyProGluArgAPro-----GluGlyPro 508
Db 1558 GGTCTGAGCGAGCCAGAGCCGCCACCACTCCAGGAGGAGGTGTGAGGGCCCC 1617
QY 509 GlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyrProGlyAlaPro 528
Db 1618 GGGGCTCAAGTACCACTCTCCAGAGGCCACAGACGGCTTCTGCTGGCAGGTGACACC 1677
QY 529 SerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAlaArgAla 546
Db 1678 AGCACAGAGCATCAGGCCAGACAGGGGCCCCAGGAGGCCACCACTGCCACTGCC 1731

RESULT 15
AX239571
LOCUS AX239571 4722 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 14 from Patent WO0164874.
ACCESSION AX239571
VERSION AX239571.1 GI:15797253

KEYWORDS Cryptolagus cuniculus (rabbit)
SOURCE Cryptolagus cuniculus
ORGANISM Cryptolagus cuniculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
REFERENCE 1
AUTHORS Lees, A.M., Lees, R.S., Law, S.W. and Arjona, A.A.
TITLE Low density liprotein binding proteins and their use in diagnosing
and treating atherosclerosis
JOURNAL Patent: WO 0164874-A 14 07-SEP-2001;
Boston Heart Foundation, Inc. (US)

FEATURES
Location/Qualifiers
1..4722
/organism="Oryctolagus cuniculus"
/mol_type="unassigned DNA"
/db_xref="taxon:9986"
61..1734
/note="unnamed protein product"

CDS
1..1734
/codon_start=1
/protein_id="CAC88294.1"
/db_xref="GI:15797254"
/db_xref="EMBL:CAC88294"
/translation="MKNDQKNGAKQNPSSPQPEAGAGAGGAGGPRPAPAREAE
GASSQAPGEGAQAKTQFALQDVSELSQLEDDILSYCVNQQAPGEDGVQGE
PEPEDAEKRAYARNGEPEPTVANGEXETSAPEGTBEIRTSVSDVDRHRPQ
EKKAQGLKGTLLMQTINTLTPPEKLAALCKYAELEEHNSQKMLQKQS
QVQKHLRGHSHKAILARSKLSICRELORHNSLKEEGVQVAREEERKEVTS
FOWTLNDLOLOMHNERSKLROENMELAEKLIROYELREHIDKVPKHQLOQ
QVQDLQAQOMLEAEERHOREKDFLLKEAVESORMCELMKQOETHLKOOLALYE
KFEETQNTUSKSEVFTTFKQEMKTKIKLERETMYSRWESSNKALLEAEK
TLRDELEGLQVKIORLEKLCRALOTERNDLNKRVODLSAGGQGVSDSGPFRPEPA
TT"SKQGVGPGAQVFNPSRATDASCCAGAPSTEASGOTGPQEPPTIATA"

ORIGIN

Alignment Scores:
Pred. No.: 5..68-112 Length: 4722
Score: 2552.50 Matches: 508
Percent Similarity: 93.19% Conservative: 12
Best Local Similarity: 91.04% Mismatches: 25
Query Match: 91.65% Indels: 13
DB: 6 Gaps: 3

US-10-023-523-44 (1-546) x AX239571 (1-4722)

QY 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerSerPro 20
Db 61 ATGAAGAATCAAGACAAAAGAACGGGGCTGCCAACAACAGCCCAACCCAAAAGCAGCCG 120
QY 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaProAla 40
Db 121 GGCACCCCGAAGCAGGAGCGGAGGAGCCAGCGGGCGGCCCGCGCGGCCCGCCCGCC 180
QY 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
Db 181 CGAGAGCCGGAAGGT---GCCACGACGACGCTCCCGGGAGCGGAGGGGGGTCAAGCC 237

QY 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
Db 238 AAAACATGCTAGCCTGGGGCGCTCTGTGTGTCTCTGAGGAGCTGAGCCGCCAGTTGGAA 297
QY 81 AspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAla 100
Db 298 GACATATCTAGTACATCTGTGTGGACAACAACCGGGGGGCCCGGGTGGAGTGGGTC 357
QY 101 GlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsn 120
Db 358 CAGGGTGGAGCCCTGAACCTGAAGATGACAGAAAGTCTCGCGCCCTATGTGGCAAGAAT 417
QY 121 GlyGluProGluPro---ThrProValValAsnGlyGluLysGluProSerLysGlyAsp 139
Db 418 GGGGAGCCGGAGCGGGCACCCAGTAGTCAATGGCGAGAGGAGACCTCCAAGGCAGAG 477
QY 140 ProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAspHisArgArgPro 159
Db 478 CCGGGCAGCGAGAGATCCGGACGAGCGATGAGTGGAGACCGGAGACCAACCGGAGGCCA 537
QY 160 GlnGlyLysLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeu 179
Db 538 CAGAAAGAAAGAGAGCCAGAGGTCTGGGAAAGAGATCACGCTGCTGATGCAGACACTG 597
QY 180 AsnThrLeuSerThrProGluGluLysLeuAlaAlaLeuCysLysLysTyrAlaGluLeu 199
Db 598 AACACGCTGAGCACCCCGAGAGGAGAGTCTGGCGCTCTGTCAAGAAGTATGCGGAATC 657
QY 200 LeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysGlnSerGln 219
Db 658 CTCGAGGAGCACCGGAACCTCGCAGAGAGATGAAGCTGCTGCAGAAAGAGAGAGCCAG 717
QY 220 LeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSer 239
Db 718 CTGCTGAGGAGAGGAGGACCACTCGGTGGCGAGCACAGCAAGGCATCTCTGGCCGCGC 777
QY 240 LysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGly 259
Db 778 AAGCTCAGAGAGCTGTGTCGGGAGCTGCAGCGGCACAAACCGCTCGCTCAAGGAAGAGGT 837
QY 260 ValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnVal 279
Db 838 GTGAGCGAGCCCGAGAGGAGGAGGAGAGCGCAAGAGGTGAGCTCACACTTCCAGATG 897
QY 280 ThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArg 299
Db 898 ACGCTCAACGACATTCAGCTGAGTGGAGCAGCACCAACGAGCGCAACTCCAAGCTCGC 957
QY 300 GlnGluAsnMetGluLeuAlaGluArgLysLysLeuIleGluGlnTyrGluLeuArg 319
Db 958 CAGGAGAACATGGAGCTGGCGAGCGGTCAAGAAAGCTGATTGACAGTACGAGCTGCGA 1017
QY 320 GluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAla 339
Db 1018 GAAGAGCACATCGACAAAGTCTTCAAAACACAAAGGATCTGCAGAGCAGAGCTGGTGGAGCC 1077
QY 340 LysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGlnArgGluLys 359
Db 1078 AAGCTCCAGCAGCGCCAGGAGATGCTGAAGAGGAGGAGGAGCGGCGCACCGGGAGAG 1137
QY 360 AspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGln 379
Db 1138 GACTTCTCTGAGAGGAGCGCTGGAGTCCCAGAGGATGTGCGAGCTGATGAAGCAACAG 1197
QY 380 GluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsn 399
Db 1198 GAGACCCACCTGAAGCAGCAGCTTGCCTATACACAGAGAGATTTGAGGAGTTCCAGAAC 1257
QY 400 ThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThr 419
Db 1258 ACTCTTTTCAAAGCAGCGAGGTGTTTACCACATTTCAAACAGGAAATGGAAGATGACA 1317

QY 420 LysLysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArgTrpGluSerSer 439
 Db 1318 AAGAGATCAAGAGCTGGGAAAGAGACACCATGTACCTGTTCCCGTGGGAGAGCAGC 1377
 QY 440 AsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGluLeuGluGly 459
 Db 1378 AACAAAGCCCTGCTTGGATGGCTGAGGAGAAACACATCCGGGACAAAGAGCTGGAGGC 1437
 QY 460 LeuGlnValLysIleGlnArgLeuGluLysLeuLysCysArgAlaLeuGlnThrGluArgAsn 479
 Db 1438 CTCAGGTGAATCATCAGCGGCTGGAGAGCTGTGCCGGGCACTGCAGACAGAGCGCAAT 1497
 QY 480 AspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSer 499
 Db 1498 GACCTGAACAGAGGCTGAGGAGCTGAGTGGCGGTGGCCAGGGCCCTCTCCACAGC 1557
 QY 500 GlyProGluArgArgPro-----GluglyPro 508
 Db 1558 GPTCCTGAGCGGAGGCGCAGAGCCCGCCACCATCTCAAGGAGCAGGGTGTCCAGGGGCCCC 1617
 QY 509 GlyAlaGlnAlaProSerProArgValThrGluAlaProCysTyrProGlyAlaPro 528
 Db 1618 GGGGCTCACTACCACTCTCCAGGGCCACAGAGCTTCTGCTGCCAGGTGCACCC 1677
 QY 529 SerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAlaArgAla 546
 Db 1678 AGCACAGAGCATCAGCCAGACAGAGGCGCCAGGAGCCACCATCTGCCACTGCC 1731

RESULT 16

BD056447 4722 bp DNA linear PAT 27-AUG-2002
 LOCUS Novel low density lipoprotein binding proteins and their use in
 DEFINITION diagnosing and treating atherosclerosis.
 BD056447
 ACCESSION BD056447.1 GI:22602053
 VERSION JP 2001506983-A/5.
 KEYWORDS Aequorea victoria
 SOURCE Aequorea victoria
 ORGANISM Aequorea victoria
 Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;
 REFERENCE 1 (bases 1 to 4722)
 AUTHORS Lees, A.M., Lees, R.S., Law, S.W. and Arjona, A.A.
 TITLE Novel low density lipoprotein binding proteins and their use in
 JOURNAL diagnosing and treating atherosclerosis
 PATENT: JP 2001506983-A 5 29-MAY-2001;
 BOSTON HEART FOUNDATION INC
 COMMENT PN JP 2001506983-A/5
 PD 29-MAY-2001
 PF 26-NOV-1997 JP 1998524870
 PR 27-NOV-1996 US 60/031930, 03-JUN-1997 US 60/048547 PI
 ANN M LEES, ROBERT S LEES, SIMON W LAW, ANIBAL A ARJONA PC
 A61K38/04, A61K38/17, A61K39/00, A61K48/00, A61K49/00, A61K51/08, PC
 C07H21/00,
 PC C07K7/00, C07K14/705, C12N15/12, C12Q1/02, C12Q1/68, G01N33/566 CC
 Strandedness: Single;
 CC Topology: Linear;
 FH Key Location/Qualifiers.

FEATURES

source
 1..4722
 /organism="Aequorea victoria"
 /mol_type="genomic DNA"
 /db_xref="taxon:6100"

ORIGIN

Alignment Scores:

Pred. No.: 5.6e-112 Length: 4722
 Score: 2552.50 Matches: 508
 Percent Similarity: 93.19% Conservative: 12
 Best Local Similarity: 91.04% Mismatches: 25
 Query Match: 91.65% Indels: 13
 DB: 6 Gaps: 3

US-10-023-523-44 (1-546) x BD056447 (1-4722)

QY 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerSerPro 20
 Db 61 ATGAAGATCAAGACAAAAGAACGGGCTGCCAAACAGCCCAACCCCAAAAGCAGCCCG 120
 QY 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnLysArgProSerGlnAlaAlaProAla 40
 Db 121 GGACAGCGGAGCAGAGCGGAGGAGCCAGGGGCGGCCCGCGCCGCGGCCCGCC 180
 QY 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
 Db 181 CGAAGCCGCAAGCT---GCCAGCAGCCAGGCTCCCGGAGGCGGAGGGGGCTCAAGCC 237
 QY 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluLysLeuSerArgGlnLeuGlu 80
 Db 238 AAAAATCTCTCAGCTGGCGGCTCTGTGATGCTCTGAGGAGCTGAGCCGCGCAGTTGGAA 237
 QY 81 AspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAla 100
 Db 298 GACATCTCAGTACATACTGTGTGGACAACAACAGGGGCGGCCCGGTGAGGATGGGTC 357
 QY 101 GlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsn 120
 Db 358 CAGGGTGAGCCCTCAACCTGAAGATGCAGAGAAGTCTCGCGCTATGTGGCAAGAAAT 417
 QY 121 GlyGluProGluPro---ThrProValValAsnGlyGluLysGluProSerLysGlyAsp 139
 Db 418 GGGGAGCGGAGCGGGGACCCCGAGTAGTCAATGGCGAGAGGAGACCTCCAAAGGAGAG 477
 QY 140 ProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAspHisArgArgPro 159
 Db 478 CCGGGCAGCGAAGAGATCCGGACGAGATAGGTGGAGACGAGACGACACCGAGGCCCA 537
 QY 160 GlnGluLysLysLysAlaLysGlyLeuGlyLysGluLeuThrLeuLeuMetGlnThrLeu 179
 Db 538 CAGAAAAGAAAGAGGCGCAAGGGTCTGGGAAAGAGATCACGCTGTGTGATGAGACACTG 597
 QY 180 AsnThrLeuSerThrProGluGluLysLeuAlaLeuLysLysLysTyrAlaGluLeu 199
 Db 598 AACACGCTGAGCACCACAGAGGAGAGCTGGCGGCTCTGTGCAAGAAAGTATCGGAACTG 657
 QY 200 LeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGln 219
 Db 658 CTGAGAGACACCGGAATCTCGAAGAGCAGATGAAGCTCTCGAAGAAAGCAGAGCCAG 717
 QY 220 LeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSer 239
 Db 718 CTGCTGCGAGGAGAGGACCACTGCTGCTGCGAGCAGCAGCAAGGCCATCTCTGCGCCGAGC 777
 QY 240 LysLeuGluSerLeuLysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGly 259
 Db 778 AAGCTCGAGAGCCTGTGCGGGAGCTGTCAGCGGCGCAACCCGCTCGCTCAAGAAAGAGT 837
 QY 260 ValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnVal 279
 Db 838 GTCCAGCGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 897
 QY 280 ThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArg 299
 Db 898 AGCTCAACGACATTCAGCTGCAGATGGAGCAGCAGCAACAGCGCAACTCCAAAGTGGCG 957
 QY 300 GlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuLeuGlnTyrGluLeuArg 319
 Db 958 CAGGAGAACATGGAGCTGGCGGCGGCTCAAGAAAGCTGATTGAGCAGTACAGCTGGGA 1017
 QY 320 GluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeuValAspAla 339
 Db 1018 GAAGACCATCGACAAAGCTTTCAACACAGAGGATCTGCACACAGCAGCTGCTGGACGCC 1077
 QY 340 LysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluGluArgHisGlnArgGluLys 359
 Db 1078 AAGCTCCAGCAGGCGCCAGGAGATGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1137

QY	360	AspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGln	379
Db	1138	GACTTCTCTCTGAAGAGCGCGTGGAGTCCACAGGATGTCGAGCTGATGAAGCAACAG	1197
QY	380	GluThrHisLeuLysGlnGlnLeuAlaLeuThrThrGluLysPheGluGluPheGlnAan	399
Db	1198	GAGACCCACTGAAGCAGCAGCTGGCTTATACACAGAGAGTTTGAGGATTCAGAAC	1257
QY	400	ThrLeuSerLysSerGluValPheThrThrPheLysGlnGluMetGluLysMetThr	419
Db	1258	ACTCTTTTCAAAACGACGCGAGTGTTCACACATTCAAAACAGAAATGGAAGAATGACA	1317
QY	420	LysLysLysLysLysGluLysGluThrThrMetThrArgSerArgTrpGluSerSer	439
Db	1318	AAGAAATCAAGAGCTGGGAAAGAGACACCATGTACGTTCCCGTGGGAGCAGC	1377
QY	440	AsnLysAlaLeuLeuGluMetAlaGluLysThrValArgAspLysGluLeuGluGly	459
Db	1378	AACAAGGCCCTGCTGAGATGGCTGAGGAGAAACACTCCGGGACAAAGAGCTGGAAGC	1437
QY	460	LeuGlnValLysLysGlnArgGluLysLeuCysArgAlaLeuGlnThrGluArgAan	479
Db	1438	CTCAGTGAATAATCCAGCGCTGGGAGAGCTGTGCCGGCACTGCAGACAGAGCGCAAT	1497
QY	480	AspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSer	499
Db	1498	GACCTGAACAGAGGGTGCAGGACTGTAGTCCGGTGGCAGGCGCCCTCTCCGACAC	1557
QY	500	GlyProGluArgArgPro-----GluGlyPro 508	
Db	1558	GGTCTGAGCGAGGCGCAGAGCGCGCCACCCACCTCCAAAGAGCAGGCTGTCAGGGCCCC	1617
QY	509	GlyAlaGlnAlaProSerProArgValThrGluAlaProCysTrpProGlyAlaPro	528
Db	1618	GGGGCTCAAGTACCACTCTCAAGGGCCACAGAGCTTCTGCTGCCAGGTGCACCC	1677
QY	529	SerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAlaArgAla	546
Db	1678	AGCACAGAGGCATCAGCCACAGACAGGGCGCCAGAGCCACCACTCCCACTGCC	1731
RESULT 17			
BC060227			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
1 (bases 1 to 4702)			
Strausberg,R., Feingold,E.A., Grouse,L.H., Derge,J.G., Klusner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zesberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullaby,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzyzanski,M.I., Skalska,U., Smallos,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.			
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences			
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)			

12477932	2 (bases 1 to 4702)	12477932	2 (bases 1 to 4702)
Strausberg,R.			
Direct Submission			
Submitted (20-OCT-2003) National Institutes of Health, Mammalian			
Gene Collection (MGC), Cancer Genomics Office, National Cancer			
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,			
USA			
NIH-MGC Project URL: http://mgc.nci.nih.gov			
Contact: MGC help desk			
Email: cgapbs-remail.nih.gov			
Tissue Procurement: Dr. James Lin, University of Iowa			
cDNA Library Preparation: M. Bento Soares, University of Iowa			
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant.			
Web site: http://genome.uiowa.edu			
Contact: bento-soares@uiowa.edu ; tom-casavant@uiowa.edu			
Bonaldo,M.F., Akabogu,I., Bair,T., Bair,J., Crouch,K., Davis,A., Fishler,K., Keppel,C., Kucaba,T., Lebeck,M., Melo,A., Schaefer,K., Scheetz,T., Smith,C., Snir,E., Tack,D., Trout,K., Walters,J., Casavant,T., Soares,M.B.			
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov			
Series: Plate: Row: Column: 0.			
Location/Qualifiers			
1..4702			
/organism="Mus musculus"			
/mol_type="mRNA"			
/strain="C57BL/6"			
/db_xref="taxon:10090"			
/clone="MGC:77972 IMAGE:30357541"			
/tissue_type="Brain"			
/clone_lib="NIH_SWAP_GHO"			
/lab_host="DH10B"			
/note="vector: pYX-ASC"			
220..1884			
/codon_start=1			
/product="Unknown (protein for MGC:77972)"			
/protein_id="AAH60227.1"			
/db_xref="GI:38614271"			
/translation="MKQDKKNGPKHNSKSGPQREAGPEGAGHPRPTAPGAAB			
GSTQAPKGTGAKAQAQPCALCDVSELSQLLEDILSTYCVNNQGGPAEAGAAGE			
PTEDPTKSRVTAARNGEPGIPVNVGKETSKEGPGTETETASDEVGDHRHPQ			
EKKAKGLKEITLLMOTLNTLSTPEEKALCKKVAELLEHNRSOKMKLLKKOS			
OLVQEKDHLSEHSKAVLBSKLELCRELQNRSLKESGVORAREEKKRKEVTS			
FQVLNDICIQMECHNENSKLROENNELAERLKLIEQELREHIDKVKKDLQ			
QLVDKLCQAQEMLKAEERHQRKEKLELKEAVESQRCMLKQOETHLKQQLALYTE			
KFEFQTLKSSSEVFTTFKQEMKMTKKIKLEKETTMYRSWESSNKALLEMAEB			
TVRDELEGLQVKIQRIEKLQALQTERNDLNKRVQDLTAGGITDGSRRPEATTAS			
KEQVESFGAQPASSPRATDAPCCGAPSTGTAGTGTGPTPTATA"			
ORIGIN			
Alignment Scores:			
Pred. No.: 2.02e-107 Length: 4702			
Score: 2456.00 Matches: 494			
Percent Similarity: 90.32% Conservative: 10			
Best Local Similarity: 88.53% Mismatches: 38			
Query Match: 88.19% Indels: 16			
DB: 10 Gaps: 4			
US-10-023-523-44 (1-546) x BC060227 (1-4702)			
QY	1	MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerSerPro	20
Db	220	ATGAAGAACCAAGACAAAGAAATGGCGCTCCCAACACTCCAACTCGAAGGCGAGCCG	279
QY	21	GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla	40
Db	280	GGGCAACGGGAAGCAGGAGCCGAGGAGCCCATGACGACCCACAGACAGCTCTCC	339
QY	41	ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla	60


```

Db      340  GCGGAAGCCGAAGT---TCCAAAGCCAGGCTCCCTGGGAAGACCCGAGGGGGCTCAGCT 396
QY      61  ArgThrAlaGlnSerGlyAlaLeuA--GAspValSerGluGluLeuSerArgGlnLeuGlu 80
Db      397  AAGCAGCTCAGCCTGGGGCCCTCTGTGACGCTCTCTGAGAGCTGAGCCGCGCAGTTGGAA 456
QY      81  AspileuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAla 100
Db      457  GACATCCTGAGTACATCTGTGTGGCAACAATTCAGGGAGGCGCCGCTGAGGAGGAGCA 516
QY     101  GlnGlyLysProAlaGluProGluAspAlaGlnLysSerArgThrTyrValAlaArgAsn 120
Db      517  CAGGTTGAGCCCTGAGCCCGGAAGACACCGAGAGCTCCGAACTATGCAGCCAGGAT 576
QY     121  GlyLysProGluPro---ThrProValValAsnGlyGlyLysGluProSerLysGlyAsp 139
Db      577  GGGGAGCCTGAACAGGCATTCAGTCTGTCACCGCGGAGAGGAGACCTCTAAGGGAGAG 636
QY     140  ProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAspHisArgArgPro 159
Db      637  CCTGGAAACAGAGAGATCCGAGGAGTGATGAAGTTGGAGACCGACCACTCGAGGCCA 696
QY     160  GlnGlyLysLysLysAlaLysGlyLysGluLysGluLeuLeuMetGlnThrLeu 179
Db      697  CAGGAGAAGAAAGCAGAGGCTAGGGAAGAGATCACTCTGCTGATGCAGACACTG 756
QY     180  AsnThrLeuSerThrProGluGluLysLeuAlaLeuLysLysLysLysLysLysLys 199
Db      757  AACACGCTGATACCCAGAGGAGAGCTGGCTGCTGTGCAAGAAAGTATGCTGAGCTG 816
QY     200  LeuGluLysHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGln 219
Db      817  CTGGAAGAGCATCGGAACCTCGAGAGCAGATGAAGCTCTGCGAGAAGAGCAGGCCAG 876
QY     220  LeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSer 239
Db      877  CTCGTGCAGGAGAAGACCATCTGCGAGGGGAACACAGCAAGGCTGCTGCGCCGAGC 936
QY     240  LysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGluGly 259
Db      937  AAGCTTGAGAGTCTGTGCGGGAGCTGCAACGGCACACACCGTCCCTGAGGAGAGAGC 996
QY     260  ValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnVal 279
Db      997  GTGCAGCGACCCGCTGAGGAGGAGGAGAGCGCAAGAGTGAATCTCACATCTCCAGGTG 1056
QY     280  ThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArg 299
Db     1057  ACATGAATGACATTCAGCTGAGTGGAACATGAGCAATGAGGAAACTCCAGCTGCGC 1116
QY     300  GlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTyrGluLeuArg 319
Db     1117  CAGGAGATATGAGCTAGCCGAGAGGCTCAAGAGTGTATCGAGCAATACGAGCTTCGT 1176
QY     320  GluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAla 339
Db     1177  GAGGAGCATATCGACAAAGTCTTCAACATAGGACCTTCAGCAGCAGCAGTGTGGACGCC 1236
QY     340  LysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluGluArgHisGlnArgGluLys 359
Db     1237  AGCTCCAGCAGGCCCGGAGAGTGTGAGAGGCGCAGAGGCGGACCCAGCAGAGAGAG 1296
QY     360  AspPheLeuLeuLysGluAlaValAlaGluSerGlnArgMetCysGluLeuMetLysGlnGln 379
Db     1297  GAGTTTCTCTGAGGAAGCGGTGGAGTCCAGAGGATGTGCGAGCTGTATGAGCAGCAG 1356
QY     380  GluThrHisLysLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsn 399
Db     1357  GAGACCCACCTTAAGCAGCAGCTCGCCCTGTACCGGAGAGTTTGGAGAGTTCCAGAAC 1416
QY     400  ThrLeuSerLysSerSerGluValPheThrPheLysGlnGluMetGluLysMetThr 419

```

```

Db      1417  ACATTTTCCAAAAGCAGTGAAGTGTTCACCACGCTTCAAACAGGAGATGGAAGAAATGACA 1476
QY     420  LysLysLysLysLysLysLeuGluLysGluThrThrMetTyrArgSerArgTTPGluSerSer 439
Db     1477  AAGAAAGATCAAGAAGCTGGGAAGAAACACCCATGTATCGATCCCGTGGGAAGCAGC 1536
QY     440  AsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGluLeuGluGly 459
Db     1537  AACAAAGGCTCTGCTGGAGATGGCTGAGAGAAAACCTCCGGGACAAAGAGCTGGAGGC 1596
QY     460  LeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsn 479
Db     1597  CTGCAAGTGAATAATCCAGCGCTGAGAAAGCTGTGCGAGCACTGCGACAGCCGAGCAAT 1656
QY     480  AspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSer 499
Db     1657  GACCTCAACAGAGGGTACAGACCTGACTGACAGGGGC-----ATCATGTGACATT 1707
QY     500  GlyProGluArgArgPro-----GluGlyPro 508
Db     1708  GGCTCTGAGCGGAGGCGCAGAGGCCACCATCGCTCCAAAGGAACAAGGGTTGAAAGTCT 1767
QY     509  GlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyrProGlyAlaPro 528
Db     1768  GGGGCTCAACAGCAGCTCTCCAAAGGGCCACAGACGCTCTCTGCTCGGAGAGCCCCG 1827
QY     529  SerThrGluAlaSerGlyGlnThrGlyProGlnProThrSerAlaArgAla 546
Db     1828  AGCAGAGAAACAGAGCCAGACAGAGGCGCTGGAGAGCCACCCCTGCGCACTGCC 1981

RESULT 18
AX578043 1793 bp DNA linear PAT 08-JAN-2003
LOCUS AX578043
DEFINITION Sequence 165 from Patent WO02081745.
ACCESSION AX578043
VERSION AX578043.1 GI:27647251
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Garcia, T., roman Roman, S., Baron, R., Call, K., Theillhaber, J.,
Connolly, F., Jackson, A., Bushnell, S. E. and Rawadi, G.
TITLE Genes involved in osteogenesis, and methods of use
JOURNAL Patent: WO 02081745-A 165 17-OCT-2002;
Aventis Pharma S.A. (FR)
FEATURES
Location/Qualifiers
1..1793
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Homo sapiens interleukin 14 (IL14), mRNA"

ORIGIN

Alignment Scores:
Pred. No.: 2,09e-73 Length: 1793
Score: 1727.00 Matches: 342
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 62.01% Indels: 0
DB: Gaps: 0

US-10-023-523-44 (1-546) x AX578043 (1-1793)
QY     205  AsnSerGlnLysGlnMetLysLeuGlnLysGlnSerGlnLeuValGlnGluLys 224
Db     1791  AATTCACAGAAGCAGATGAGCTCTACAGAAAAGCAGAGCCAGCTGGTGCAGAGAG 1732
QY     225  AspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeu 244
Db     1731  GACCACCTGCGCGGTGAGCACAGCAAGCGCTCTGCGCCCGCAGCAAGCTTGAGAGCTA 1672

```


	Qy	245	CysArgGluLeuGlnArgHisAsnArgSerIeuLysGluGlyValGlnArgAlaArg	264
	Dd	1671	TGCCTGAGTGGACGGCGCAACCCGTCCTCCTCAAGAAGAAGGTGTGCAGCGGCCCGG	1612
	Qy	265	GluGluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspile	284
	Dd	1611	GAGGAGGAGGAAGCGCAAGGAGGTGACCTCGCACCTCCAGGTGACACTGAATGACATT	1552
	Qy	285	GlnLeuGlnMetGluGlnHisAsnGluUArgAsnSerLysLeuArgGlnGlnAsnMetGlu	304
	Dd	1551	CAGCTGCAGATGGAACAACGCAATAGCGCAACTCCAAGCTGGCGCCAAAGAACATGGAG	1492
	Qy	305	LeuAlaGluArgLeuLysLysLeuIleGlnTyrgluLeuArgGluGluHisIleAasp	324
	Dd	1491	CTGGCTGAGAGGCTCAAGAGACTGATTGAGCATGTAGCTGGCGAGGAGCATATCGAC	1432
	Qy	325	LysValPheLysHisLysAspLeuGlnGlnLeuValAspAlatLysLeuGlnGlnAla	344
	Dd	1431	AAAAGTTCTCAAACACAAAGGACCCTACACAGCAGCTGCTGGTAGTCCAAAGCTCCAGCAGGCC	1372
	Qy	345	GlnGluMetLeuLysGluAlaGluGluArgHisGlnArgGluLysAspPheLeuLys	364
	Dd	1371	CAGAGATGCTTAAGGAGGCGAAGAGCGGCACCAGCGGGAGAAAGGATTTTTCTCTGAAA	1312
	Qy	365	GluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLeuLys	384
	Dd	1311	GAGCAGTAGAGTCCCAGAGAGTGTTGTGAGCTGATGAAGCAGCAAGAGAGACCCACCTGAAG	1252
	Qy	385	GlnGluLeuAlaLeuTyrrThrGluLysPheGluCluPheGlnAsnThrLeuSerLysSer	404
	Dd	1251	CAACAGCTTCCCTTATACAGAGAAAGTTTGAGGAGTTCCAGAAACACACTTCCAAAGC	1192
	Qy	405	SerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysIleLysLys	424
	Dd	1191	AGCGAGTATTCCACCACATTCAAGCAGGAGATGGAAAAAGATGACTAAGAAGATCAAGAAG	1132
	Qy	425	LeuGluLysGluThrThrMetTyrrArgSerArgTrpGluSerSerAsnLysAlaLeuLeu	444
	Dd	1131	CTGGAGAAAGAAACCCACTGTACCGGTCCGGTGGGAGAGCAGCAACAGGCGCTGCTT	1072
	Qy	445	GluMetAlaGluGluLysThrValArgAspLysGluLeuGluGlyLeuGlnValLysIle	464
	Dd	1071	GAGATGCTCAGGAGAAACAGTCCCGGATAAAGAACTGGAGGGCTCGCAGGTAAAAATC	1012
	Qy	465	GluArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArg	484
	Dd	1011	CAACGGCTGGAGAAGCTGTCCGGGCACTGCAGACAGGGCGAATGACCTGAAACAAGAGG	952
	Qy	485	ValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSerGlyProGluArgArg	504
	Dd	951	GTACAGGACCTGAGTGTGTGTGGCCAGGGCTCCCTCACTGACAGTGGCCCTGAGAGGAGG	892
	Qy	505	ProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyr	524
	Dd	891	CCAGAGGGGCTGGGGCTCAAGCACCAGCTCCCCCAGGGTCAAGAGAGGGCCTTGTGTAC	832
	Qy	525	ProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAla	544
	Dd	831	CCAGGACCCGAGCACAGAAGCATACAGCCAGACTGGGGCTCAAGAGGCCACCTCCGCC	772
	Qy	545	ArgAla	546
	Dd	771	AGGGCC	766

RESULT 19	HUMHMBCEGF/c	1793 bp	linear	PRI 25-MAY-1995
LOCUS	HUMHMBCEGF	weight	B cell growth factor mRNA	
DEFINITION	Human high molecular weight B cell growth factor mRNA sequence.			
ACCESSION	L15344			
VERSION	L15344.1	GI:832913		
KEYWORDS	B-cell growth factor; interleukin 14.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			

REFERENCE AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (bases 1 to 1793) Ambrus J.L. Jr., Pipkin, J., Joseph, A., Xu, C., Blumenthal, D., Tamayo, A., Claypool, K., McCourt, D., Srikiatchoorn, A. and Ford, R. J.
TITLE	Identification of a cDNA for a human high-molecular-weight B-cell growth factor
JOURNAL MEDLINE	Proc. Natl. Acad. Sci. U.S.A. 90 (13), 6330-6334 (1993)
PUBLISHED	93317675
REFERENCE	8327514
AUTHORS	2 (bases 1 to 1793) Ambrus J.L. Jr.
TITLE	Direct Submission
JOURNAL	Submitted (03-MAY-1993) Julian L. Ambrus, Jr., Department of Medicine, Washington University School of Medicine, St. Louis, MO 63110, USA
COMMENT	On May 25, 1995 this sequence version replaced gi:347805. Original source text: Homo sapiens cDNA to mRNA.

FEATURES
source

ORIGIN

Alignment Scores:		
Pred. No.:	1.46e-71	Length:
Score:	1688.00	Matches:
Percent Similarity:	99.12%	Conservative:
Best Local Similarity:	99.12%	Mismatches:
Query Match:	60.61%	Indels:
DB:	9	Gaps:
US-10-023-523-44 (1-546)	x HUMHMBECCF	(1-1793)

Qy	205	AsnSerGlnLysGlnMetLysLeuLeuGlnLysGlnSerGlnLeuValGlnGluLys	224
Db	1791	AATTACAGAAGCAGATGAAGCTCCTACAGAAAAGCAGAGCCAGCTGGTGCAGAGAAG	1732
Qy	225	AspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeu	244
Db	1731	GACCACCTGGCGGCTGAGCACACAGCC-GTCTGCGCCGCAAGACTTGAGAGCCTA	1673
Qy	245	CysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGlyValGlnArgAlaArg	266
Db	1672	TGCCGTGAGCTGACGGGCACAACCGCTCCCTCAAGGAAGAAGGTGTGCAGCGGCGCCGG	1613
Qy	265	GluGluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIle	284
Db	1612	GAGGAGGAGGAGAAGCGCAAGGAGGTGACCTCCGCACTTCGAGGTGACACTGAATGACATT	1553
Qy	285	GlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGlu	304
Db	1552	CAGCTGCAGATGGAACAGACACAATGAGCGCAACTCCAAAGCTGCGCCAAGAGAACAATGGAG	1493
Qy	305	LeuAlaGluArgLeuLysLysLeuIleGluGlnTyrCluLeuArgGluGluHisIleAsp	324
Db	1492	CTGGCTGAGAGGCTCAAGAAGCTGATTGACAGATGAGCTGCGGAGGAGCATATCGAC	1433
Qy	325	LysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAla	344
Db	1432	AAAGTCTTCAAAACACAAGGACCTTACAACAGCACCTGGTGGATGCCAAGCTCCACGACGGCC	1373
Qy	345	GlnGluWerLeuLysGluAlaGluArgHisGlnArgCluLysAspPheLeuLeuLys	364
Db	1372	CAGGAGATGCTTAAGGAGGCCAAGAAGCGGCCACCGGGAGAAAGGNTTTTCTCTGAAA	1313
Qy	365	GluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLeuLys	384
Db	1312	GAGCGAGTAGTGCCACAGAGATGTGTGACTGATGAAGCAGACAAGACGCCACCTGAAG	1253

QY 385 GlnGlnLeuAlaLeuTyThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSer 404
 DB 1252 CAACAGCTTGCCCTATACACAGAGAGTTTGAGGAGTTCCAGAACACACTTCCAAAGGC 1193
 QY 405 SerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLys 424
 DB 1192 ACGGAGGTATTCACACATTCACGAGGAGATGGAAAGATGACTTAAGAGATCAAGAG 1133
 QY 425 LeuGluLysGluThrThrMetTyrArgSerArgTTPGluSerSerAsnLysAlaLeuLeu 444
 DB 1132 CTGGAGAAAGAAACACCATGATGACCGGTCCCGGTGGAGAGCAGCAACAGCCCTGCTT 1073
 QY 445 GluMetAlaGluGluLysThrValArgAspLysGluLeuGluGluLysValLysLys 464
 DB 1072 GAGATGGCTGGAGAGAAACATCGCGGATAAAGAACTGGAGGGCTCGAGGTAAATC 1013
 QY 465 GlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArg 484
 DB 1012 CAACGGCTGGAGAGCTGTGCGCGGCACCTGCACAGAGCGCAATGACTGAACAAGAGG 953
 QY 485 ValGlnAspLeuSerAlaGlyGlnGlySerLeuThrAspSerGlyProGluArgArg 504
 DB 952 GTACAGGACCTGAGTGCTGGTGGCAGGCTCCCTCACTGACAGTGGCCCTGAGGAGG 893
 QY 505 ProGluGlyProGlyValAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyr 524
 DB 892 CCAGAGGGGCTGGGCTCAACGACCCAGCTCCCGGAGGTCACAGAGCGCTTGCTAC 833
 QY 525 ProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerLys 544
 DB 832 CCAGAGGACCCGAGCAGACAGAGCATCGGCGCAGCTGGCGCTCAAGAGCCCACTCCGCC 773
 QY 545 ArgAla 546
 DB 772 AGGCCC 767

RESULT 20
 AC108815
 LOCUS Mus musculus clone RP23-239013, WORKING DRAFT SEQUENCE, 5 unordered
 DEFINITION pieces
 AC108815
 VERSION AC108815.4 GI:28882391
 KEYWORDS HTG; HTGS; PHASE1; HTGS; DRAFT.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Birren,B., Nusbaum,C. and Lander,B.
 Mus musculus, clone RP23-239013
 Unpublished
 2 (bases 1 to 181547)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Bouckgalter,B.,
 Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
 Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
 Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
 Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
 Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
 Lander,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C.,
 Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
 McEwan,P., McKernan,K., Meldrum,J., Meneus,L., Mihova,T.,
 Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
 Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
 Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
 Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
 Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
 Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,

Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (31-JAN-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 181547)
 Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
 Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
 Boguslavsky,L., Bouckgalter,B., Camarata,J., Chang,J., Choepel,Y.,
 Collymore,A., Cook,A., Cooke,P., Corum,B., DeArrellano,K.,
 Diaz,J.S., Dodge,S., Doolley,K., Dorris,L., Erickson,J., Faro,S.,
 Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
 Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
 Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
 Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
 Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C.,
 Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
 Meldrum,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
 Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
 O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
 Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
 Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
 Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
 Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
 Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
 Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (08-MAR-2003) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Mar 8, 2003 this sequence version replaced gi:28195423.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L20929
 Center clone name: 239_0_13
 ----- Summary Statistics
 Sequencing vector: Plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator; Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 180713 bases at least Q40
 Consensus quality: 180940 bases at least Q30
 Consensus quality: 181059 bases at least Q20
 Insert size: 172000; agarose-fp
 Insert size: 181147; sum-of-ctdigs
 Quality coverage: 8.0 in Q20 bases; agarose-fp
 Quality coverage: 7.6 in Q20 bases; sum-of-ctdigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 5 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1 15598: contig of 15598 bp in length
 * 15599 15698: gap of 100 bp
 * 15699 17443: contig of 1745 bp in length
 * 17444 17543: gap of 100 bp
 * 17544 39808: contig of 2265 bp in length
 * 39808 39908: gap of 100 bp
 * 39909 76850: contig of 36942 bp in length
 * 76851 76950: gap of 100 bp
 * 76951 181547: contig of 104597 bp in length.
 *
 * Location/Qualifiers
 1. 181547

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 FEATURES
 source

```
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone_lib="RP23-239013"
/clone_lib="RP23-239013" Female Mouse BAC"
1..15598
/note="assembly_fragment"
clone_end:SP6
vector_side:left
15699..17443
/note="assembly_fragment"
17544..39808
/note="assembly_fragment"
39909..76850
/note="assembly_fragment"
76951..181547
/note="assembly_fragment"
clone_end:T7
vector_side:right"

misc_feature
misc_feature
misc_feature
misc_feature
misc_feature

ORIGIN
Alignment Scores:
Pred. No.: 1.54e-61 Length: 181547
Score: 1520.00 Matches: 415
Percent Similarity: 43.13% Conservative: 35
Best Local Similarity: 39.8% Mismatches: 92
Query Match: 54.58% Indels: 505
DB: 2 Gaps: 6

US-10-023-523-44 (1-546) x AC108815 (1-181547)
QY 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerPro 20
DB 110694 ATGAAGACCAAGACAAAGAAAGATGGCTCCCAACACTTCACTCCAGAGCAACCGG 110753
QY 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
DB 110754 GGGCAAGGGGAGCAGGAGCTGGACGAGGCCATAAAGAGCCACAGACAGACGCTCTGGG 110813
QY 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
DB 110814 GCAGAGCTGAAGGT---TCCATAGCCAGGTTCTCTGGAGAGATGAGGGTCTCGAGCT 110870
QY 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
DB 110871 AAAGCAGCGTAGACTGGGGCTCTATGTCTCTGAGGAGCTGAGCGCGCAGTTGGAA 110930
QY 81 AspileLeuSer-ThrTyrcysValAspAsnGlnGlyProGlyGluAspGlyAl 100
DB 110931 GACATCTGAGTAAATAATGTGTGCAACAATCGG-----GGAGACTCGGGCAA 110981
QY 100 aGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrcysValAlaAlaArgAs 120
DB 110982 GGA-GGTGAGCCCACTGAGCGGAGAGACATAGAGAAGTCTCTGAACCTATGCGAGCAGGAA 111040
QY 120 nGlyGluProGluPro----ThrProValValAsnGlyGluLysGluProSerLysGlyAs 139
DB 111041 TGGGAAGCCTGAATCAGGCATTCACAGTCCAGTAGTCAATGGCAAGAGGAGACCTCTAAGGGGGA 111100
QY 139 pProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHisArgGpr 159
DB 111101 CGCTGTACAGAGAGATCCGAGCAGTAGTGAGGTTGGAGACTGAGACCATCGAGGCA 111160
QY 159 oGln----- 160
DB 111161 ACAAGAAAAGAGAACTCAAGAAGGTGACTTCAGAAATCAAATAACCCGTTAAAAAAT 111220
QY 160 ----- 160
DB 111221 GGGGCTCAGAACTGACAAAGGATTTCTCAGCTGAGGATACCGAATGGCAGAGAGCACC 111280
QY 160 ----- 160
```

```
DB 111281 TGAATAAATGTTCAACATCCTTAATCATCAGGGAATGCAAAATCAAAACAACCCCTGAGAT 111340
QY 160 ----- 160
DB 111341 TCCACCTCACACCAGTCAGATGGCTAAGATCAAAATTCAGGTGACAGCAGATGCTGGC 111400
QY 160 ----- 160
DB 111401 GTGGATGTGGAGAAAGAGAAACACACTCTCCATTTGTTGGTGGATTTGAGGCTTTGTACAAC 111460
QY 160 ----- 160
DB 111461 CACTCTGGAATCAGTCTGCTGTTCTCAGAAAAATTGGACACAGTACTACTGGAGGATC 111520
QY 160 ----- 160
DB 111521 CCGCAATACCTCTCTGGGCATATATCCAGAGATGTCCCAACCGGTAAAGAGACACAT 111580
QY 160 ----- 160
DB 111581 GTCCTACTATGTTTCATAGACGCTTATTATATAGCCAGAGCTGGAAGAACCCAGAT 111640
QY 160 ----- 160
DB 111641 GCCCTCAACAGAGGAATGGATACAGAAAAATGTGTACATCTACACAATATGTACTACTGATAAGTGGATA 111700
QY 160 ----- 160
DB 111701 CAGCTATTAATAAATTTATGAATTCCTAGCCAAATGGATGGACCTGGAGGGCATCATCT 111760
QY 160 ----- 160
DB 111761 GAGTGGATTAACACATTCACAAAGGAACTACACAATATGTACTACTGATAAGTGGATA 111820
QY 160 ----- 160
DB 111821 TTAGCCAAAACCTAGGATACCCAGATATAAGATACAATTTGCTAAACACATGAAATC 111880
QY 160 ----- 160
DB 111881 AAGAAAAATGAAGCTGAAGTGGACACTATGCCCTCTCTTAGAAGTGGGAACAAAACA 111940
QY 160 ----- 160
DB 111941 CCCTTGAATGAGTTACAGAGACAAAAGTTTGGAGCTGAGATGAAGAGATGGTCCATGTAG 112000
QY 160 ----- 160
DB 112001 AGACTTCCATATCCAGGATCCACCCCATATACAGCATCCAAACGCTGACCCCATTTCT 112060
QY 160 ----- 160
DB 112061 AACTAGCAAGATTTTATTGAAGAGCCAGATGTAGTCTCTCTTGTGAGACTATGCCG 112120
QY 160 ----- 160
DB 112121 GGGCTTAGCAACACAGAAAGTGGATGCTCACAGTCAGCTAATGGATGATCAGAGGCCC 112180
QY 160 ----- 160
DB 112181 CCAATGGAGAGCTAGAGAAAGTACCCAGGAGCTAAAGGATCTCAACCCCTATAGGTG 112240
QY 160 ----- 160
DB 112241 GAACAACATTATGAACCTAACAGTACCCCGAGCTCTTGACTCTAGCTCATATGTATCA 112300
QY 160 ----- 160
DB 112301 AAAGATGGCTAGTCTGGCCATCATCTGGAAGAGAGAGCCCAATGGACACGCAAACTTTATA 112360
QY 160 ----- 160
DB 112361 TGCCCCAGTACAGGGGAACGCCAGGGGCCAAAAGGGGGAGTGGGTGGGTAGGGATTGGG 112420
```



```

QY 48 SerSerGlnAlaProArgLys-----ProGluGlyAlaGlnAlaArgThrAlaGlnSer 65
Db 130 CGCGACGCGAGCCCGCGAGAGTTGTAATTTGCAATGGAAGAAAGCTGGATTTGT 189
QY 66 GlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGluAspLeuSerThr 85
Db 190 GGGCTAGGGGTGAAGCAGATATGTTGTAATCTCTCAATCAATGATATCTTCAACAT 249
QY 86 TyrCysValAspAsnGlnGlyProGlyGluAspGlyAlaGlnGlyGluProAla 105
Db 250 CAA-----GGCTCAAAATGCTGGCACAAGATAACAAGCATTTCATTGGAAGAGATGAA 303
QY 106 GluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsnGlyGluProGluPro 125
Db 304 GCAGTGACTTTATACAGAGAACAGGAATTTGGTGAGC-----CCA 345
QY 126 ThrProValValAsnGlyGluLysGluProSerLysGlyAspProAsnThrGluGluile 145
Db 346 GCATACCTGCACCAAGAATCAAGAGAGGAATCCCTGGGGGA-----GAAGCT 393
QY 146 ArgGlnSerAspGluValGlyAspArgAspHisArgArgProGlnGluLysLysAla 165
Db 394 CAAACAGATCCCTCGATGTCAGCAAGATTCAGAGTGCACAGCAACAGCAAAAGAAAACT 453
QY 166 LysGlyLeuGlyGluLeuThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrPro 185
Db 454 -----TTAGGAAAAGAGTTTATTACTGATGCGAAGCCCTAAACACCTTTCAACCCCA 507
QY 186 GluGluLysLeuAlaLeuCysLysLysTyrAlaGluLeuLeuGluGluHisArgAsn 205
Db 508 GAGGAGAACTGCGAGCTCTCTGTAGAAATATGCTGATCTTCTGGAGAGCAGGAGT 567
QY 206 SerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAsp 225
Db 568 GTTCAGAGCAANTGAGATCTCTGAGAGAGAGCAAGCCAGATTGTGAAGAGAAAGTT 627
QY 226 HisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCys 245
Db 628 CACTTGCAGAGTGAACATAGCAAGCTATCTTGGCAAGCAAGCAAGCTAGATCTCTTTGC 687
QY 246 ArgGluLeuGlnArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGlu 265
Db 688 AGAGAACTTCAGCGTCACAATAGACGTTAAGAGGAGGAAATATATGCAGCGACGACGAG 747
QY 266 GluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGln 285
Db 748 GAAGAAGAACGACGTAAGAAGCACTGCACATTTCCAGATTACCTTAGATGAATTCAA 807
QY 286 LeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGluLeu 305
Db 808 GCCCAGCTGGAGCAGCATGACATCCACACGCCAACTCCGACAGGAAACATTGAGCTG 867
QY 306 AlaGluArgLeuLysLysLeuIleGluGlnTyrGluLeuArgGluGluHisLysAspLys 325
Db 868 GGGGAGAGACTTAAGAGACTCATCGACATGACGACTGAGGAGAGACCATTTGATTAAG 927
QY 326 ValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGln 345
Db 928 GTGTTCAAAACGTAAGAACTGCAACAGCAGCTCGTGGATGCCAACTGCAGCAAAACGACA 987
QY 346 GluMetLeuLysGluAlaGluGluArgHisGlnArgGluLysAspPheLeuLysGlu 365
Db 988 CAACCTGATAAAGAGCTGATGAAAAACATCAGAGAGAGAGAGTATTTTATTAAGAA 1047
QY 366 AlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlnThrHisLeuLysGln 385
Db 1048 GCGACAGAATCGAGCACAATAACGACAAATGAAACAGCAGGAGTACAACTAAACAG 1107
QY 386 GlnLeuAlaLeuThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSer 405
Db 1108 CAGCTTCTCTTATATGATAAGTTTGAAGAAATTCACAGACTACCATGGCAAAAGCAAT 1167
QY 406 GluValPheThrThrPheLysGlnGlnMetGluLysMetThrLysLysIleLysLysLeu 425

```

```

Db 1168 GAATGTTTCAACCTTCAGACAGGAATGAAAAGATGACAAAGAAATTAATAAACTG 1227
QY 426 GluLysGluThrThrMetTyrArgSerArgTrpGluSerSerAsnLysAlaLeuLeuGlu 445
Db 1228 GAAAAGAAACAATAATTTGGCGTACCAATGGGAAACAATAATAAAGCATTCTTGCAC 1287
QY 446 MetAlaGluLysThrValArgAspLysGluLeuGlyLeuGlnValLysIleGln 465
Db 1288 ATGCTTGAAGAGAAAACAGTCCGTGATAAGAGTACAAGGCCCTTCAATAAAACTGAA 1347
QY 466 ArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgVal 485
Db 1348 CGTTAGAGAGCTGTGACGGCTCTTCAACACAGAAAGGAATGAGCTCAATGAGAAGTG 1407
QY 486 Gln-----AspLeu 488
Db 1408 GAAGTCTCTGAAAGACGAGGTATCCATCAAAGCGCCATCAAAGCGGCAACAGGATTTA 1467
QY 489 SerAlaGlyGlnGlnGlySerLeuThr-----AspSerGlyProGluArgArgProGlu 506
Db 1468 GCAACACTGTGATGACGCGCTGTACTGCTGCTGATCTTCAAGAGAGCTGGAACACTTC 1527
QY 507 GlyPro-----GlyAlaGlnAlaProSer-SerProArgValThrGluAlaProCys 523
Db 1528 TCGAAAGAGCCCTGGGAGCGCACCTGGAGGCTGAGCCCAAGATCAGAGAGCGCTGTG 1587
QY 523 sTyrProGlyAlaPro 528
Db 1588 CAAAGCCCCCGTCCA 1603

RESULT 22
BD157034 2045 bp DNA linear PAT 17-JAN-2003
LOCUS Primer for synthesizing full-length cDNA and use thereof.
DEFINITION BD157034
ACCESSION BD157034
VERSION BD157034.1 GI:27862792
KEYWORDS JP 2002191363-A/11877.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2045)
AUTHORS Ota,T., Isoqai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002191363-A/11877 09-JUL-2002;
HELIIX RESEARCH INSTITUTE
COMMENT OS Homo sapiens (human)
PN JP 2002191363-A/11877
PD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
PI SAITO,
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI KEIICHI NAGAI, TETSUJI OTSUKI
PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/
10,
PC, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
FT CDS Location/Qualifiers
source 1..2045
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores: 1.66e-52 Length: 2045
Pred. No.: 1286.00 Matches: 284
Score:

```



```

RESULT 24
HSM807779      4625 bp      mRNA      linear      PRI 30-AUG-2003
LOCUS          Homo sapiens mRNA; cDNA DKFZp451B226 (from clone DKFZp451B226).
DEFINITION     BX647633
ACCESSION      BX647633.1  GI:34366790
VERSION        Homo sapiens (human)
KEYWORDS       Homo sapiens
SOURCE         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 4625)
AUTHORS        Koehrer,K., Beyer,A., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
               Fobo,G., Han,W., and Wiemann,S.
CONSTRM        The German Human cDNA Consortium
TITLE          Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
JOURNAL        Neuherberg, GERMANY
COMMENT        Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
               Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
               sequenced by BMFZ (Biomedical Research Center at the
               Heinrich-Heine-University, Dueseldorf/Germany) within the cDNA
               sequencing consortium of the German Genome Project. This clone
               (DKFZp451B226) is available at the RZPD in Berlin. Please contact
               the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
               Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further
               information about the clone and the sequencing project is available
               at http://mips.gsf.de/proj/cDNA/.
FEATURES       Location/Qualifiers
               1..4625
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="DKFZp451B226"
               /tissue_type="human skeletal muscle"
               /clone_lib="451 (synonym: hicci). Vector pSport1; host
               DH10B; sites NotI + SalI"
               /dev_stage="adult"
               polyA_signal 4583..4588
               polyA_site 4605
ORIGIN
Alignment Scores:
Pred. No.:      1.15e-51      Length:      4625
Score:          1276.00      Matches:     277
Percent Similarity: 61.4%      Conservative: 94
Best Local Similarity: 45.8%      Mismatches:  141
Query Match:     45.82%      Indels:      92
DB:              9           Gaps:        10

US-10-023-523-44 (1-546) x HSM807779 (1-4625)
QY      3  AsnGlnAspLysAlaAlaLysGlnSerAsnProLysSerProGlyGln 22
Db      101  ATACTCTGACACCTCTCAGCGGAAACGACAGTCAACACCT-----CCAGGTGAC 151
QY      23  ProGluAlaGlyPro-----GluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
Db      152  AGTTTCATCATTCACCCAGTCACAAATGCGCTGGAGAGGAAGATGGCCAG----- 199
QY      41  ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
Db      200  -----GATCTCCACCCAGCTCCACACACGAGAAA----- 232
QY      61  ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
Db      233  -----GAGGCAAGTGTGCACCCCGGATATCTCTGAAGAGCTGAATCGACAGCTGGAA 283
QY      81  AspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAla 100
Db      284  GACATCATTAACACTTAT-----GGGTCTGTCGCCACACGACGAGGGAAGAGGGCTCT 337

```

```

101  GlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsn 120
Db      338  GCAGGGCCAGTCAGACGCTGAGAAATGAGAAATCAGAAATCAGAAATCAGAAATCAGAAAT 388
QY      121  GlyGluProGluProThrProValValAsnGlyGluLysGluProSerLysGlyAspPro 140
Db      389  GGGGACTGTGAGGAAACAACTGAAGAGGCTGGAAGAGAAACCCGCTTCTGAGAGCCA 448
QY      141  AsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAspHisArgProGln 160
Db      449  CCCACTGTCTCAAA-----GAGCCCGTCAGCAATAAGAGCAAAA-----TTG 490
QY      161  GluLysLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsn 180
Db      491  GAAAGAAATCCTAAAGGATAGCAAGAGCAACCTGCTAATGCAAAATCTGAAC 550
QY      181  ThrLeuSerThrProGluLysLeuAlaLeuCysLysLysTyrAlaGluLeu 200
Db      551  AAGTTGCAAAACACCGGAAGAAAGTTTGAATTTTATTCAGAAAGTATCTGAATGCTG 610
QY      201  GluGluHisArgAsnSerGlnLysGlnMetLysLeuLysLysGlnSerGlnLeu 220
Db      611  GATGAACATCTGACTGAGCAAAAGAGTTAAGTCTCTCAAAGAAACAGAGTACAAAT 670
QY      221  ValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLys 240
Db      671  CAAAAGAAAGAGCAGCTTACAGGTGAACACAGCAGAGTATCTCTGCTCGAAGCAAA 730
QY      241  LeuLysSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGlyVal 260
Db      731  TTGGAGAGTCTGTGCGGGAGCTGCGAGAGACACACAGAGCTCTGAAGAAAGAGGGCTT 790
QY      261  GlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnValThr 280
Db      791  CACGGGCGCAGCTGAGGAGAGAGAGAAAGAAAGAAATCACAAGCCATTCACAGAGTACC 850
QY      281  LeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGln 300
Db      851  CTCACGGACATCCAGGGCCAGATCGAGCAGCAGAGTGAGCGAAATATGAAGTCTGTCTAG 910
QY      301  GluAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTyrGluLeuArgGlu 320
Db      911  GAGAAACACAGAGCTTCAGAGAAAGCTGAAAGCATCATCATGATGATGATGATGATGATG 970
QY      321  GluHisIleAspLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLys 340
Db      971  GAGCATCTGCACAAAATATTTAAACACAGAGAACTGCGACAGAGAGCTGTGTGATGCAAG 1030
QY      341  LeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGlnArgGluLysAsp 360
Db      1031  CTTGACAGGCCCAAGAAATGATGAGAGAGCGGAGGCGAGCGACACAAACGAGAAAGAA 1090
QY      361  PheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlu 380
Db      1091  TATTTGCTGAACACGAGCAGAGTGGAAACTTCAGCGCAAAATGCTGAAGGAGCAAGAG 1150
QY      381  ThrHisLysLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThr 400
Db      1151  ACAGTCTCTGAGGCTCAGCTCTACTCTTACTCAGGAGGTTTGAGAAATTCAGAGCACA 1210
QY      401  LeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLys 420
Db      1211  CTAACATAAAGCAACGAGGTGTTTGCCACGTTCAAAACAGGAAATGCAAAAACAATAAG 1270
QY      421  LysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArgTrpGluSerSerAsn 440
Db      1271  AAAATGAAGAGCTGAAAGAGGACACAGCCCATGGAAGAGCCCGATTTGAGAACTGTAAAC 1330
QY      441  LysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGluLeuGlyLeu 460
Db      1331  AAAGCTCTGTTGGACATGATTGAAGAGAGAACACCTGAGAGCTAAAGAAATATGAGTCTTT 1390
QY      461  GlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAsp 480

```



```

Db      1391 GTGATGAATAATCGGAGGCTAGAGAACCTCTGCCGTGCTTTTCAAGAGAGAGAAACGAA 1450
Qy      481 LeuAnlyArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSerGly 500
Db      1451 CTCACAAAATAATCAGACGCGAGAAATATCTGAAAGAGATGACCAAGTCAGACACAC 1510
Qy      501 ProGluArgProGluGly----- 507
Db      1511 TCGATGAAGAGCGCAGAGTCAACGCTCTCTGTGGATCAAGAGATTGACGAGAGAGGTT 1570
Qy      507 ----- 507
Db      1571 AATAGTGTCACAAACCCCGTGAAATCTGGCCACAGCCCTTCATGATTAATCATCATCA 1630
Qy      507 ----- 507
Db      1631 GAGTCAACCCCGCAGCAGTCCAAAGAAACCCCAACCGAAATAGGCAGTCTTCAGGAGAGT 1690
Qy      508 -----ProGlyAlaGlnAlaProSerSerPro-----ArgValThrGlu 520
Db      1691 GCTGAGCGCGTCTCAAGAGCGCAGAGCAACCCCTCTGATCCCTTCACGGGATTCAGAG 1750
Qy      521 AlaProCysTy-ProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGlu 540
Db      1751 AGTCCCTCGCTCCCTAACTCTCTCAGGCTGAGCGGAGGAGGAGGATGCTGAACT 1810
Qy      541 ProThrSerAla 544
Db      1811 CCTCCAGGCC 1822

RESULT 25
HSM803295      4640 bp      mRNA      linear      PRI 13-MAY-2003
LOCUS      Homo sapiens mRNA; cDNA DKFZp451G083 (from clone DKFZp451G083).
DEFINITION      AL831988
ACCESSION      AL831988.1  GI:21732527
VERSION
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Oeanger,A.,
Fobo,G., Han,M. and Wiemann,S.
Direct Submission
Submitted (13-MAY-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.
This clone (DKFZp451G083) is available at the RZPD in Berlin.
Please contact the RZPD, Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
Location/Qualifiers
    1..4640
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="RZPD:DKFZp451G083"
    /db_xref="taxon:9606"
    /map="6q24.1"
    /clone="DKFZp451G083"
    /tissue type="human skeletal muscle"
    /clone_lib="451 (synonym: hlcc1). Vector pSport1; host
    DH10B; sites NotI, SalI"
    /dev_stages="adult"
    1..4640
    /gene="DKFZp451G083"
    <1..2158
    /gene="DKFZp451G083"

```

```

/note="similar to 77K muscle-derived protein (Gallus
gallus)"
/codon_start=2
/product="hypothetical protein"
/protein_id="CAD91140.1"
/db_xref="GI:30722291"
/translation="LFPCCSCSDLLPSPFLSLFLSPDTPPMKMEANHSEQLS
AERQSTPPSSSLPSPSHNGLEKEDGQSDPTVPPEKEASVHPDISBELANQLIEDIN
TYGSAATAGKEGASARASEOPENASPDNEDGCEETTEAGREPVASGPPTVKPEV
SNKEKLEKILKLGKEANLNLNKLQTPREKPTDFLPKYAEILDEHRTQKLLK
LLOKYOIOKEKDOLGEGHRAILARSKLESICRELORHNKILKEALORAREBEK
KRELSHFQSTLIDIQIEQSERNNKJCQENTELAEKLSIIDQELREHLDKILF
KHRELQKLVDAKLEQAEQWKEABERHKEKYLNLNQAABWKLQAKXQELVLOA
QLTLYSGRFEFQSTLTKNVFATFKQEMDKTKTKMKLEKDTATWKAFFENKAL
LDMIEKALRAKYEYCFWMKIGLELNCALQBERNELHKIKDAETSEKDDOSQHS
DEPESNVSDQIEDAEVNSVOTAVKNLATAFMIHHPESTPHOSKETOPETGSSOE
SADAALKEPQPLIPSRDSNPPLPTPOAEAGGSDAEPSPSKASPAAGLAETQC
EGLPVGTCAQPSWKPEAEASGQAPQAPTEASIQKREADVAPACAAEHVAAMVPAC
EPGRQPPRAAEELPVGASAGPOPRNVADTNLEGVD"
4598..4603
/gene="DKFZp451G083"
4620
/gene="DKFZp451G083"

polyA_signal
polyA_site

ORIGIN
Alignment Scores:
Pred. No.:      1..15e-51      Length:      4640
Score:          1276.00      Matches:      277
Percent Similarity: 61.42%      Conservative: 94
Best Local Similarity: 45.86%      Mismatches:  141
Query Match:    45.82%      Indels:       92
DB:              9          Gaps:          10

US-10-023-523-44 (1-546) x HSM803295 (1-4640)
Qy      3 AsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerSerProGlyGln 22
Db      113 AATCACTCTGAACAGCTCTCAGCGGACGACGACACCT-----CAGGTGAC 163
Qy      23 ProGluAlaGlyPro-----GluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
Db      164 AGTTCATCATTAACCTGATCAATGCTGGAGAGAGAGATGCGCCAG----- 211
Qy      41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
Db      212 -----GATTCTCCACCCGCTGAGAGAGAGAGATGCGCCAG----- 244
Qy      61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
Db      245 -----GAGCGAAGTGTGACCCCGATATCTCTGAAGAGCTGAATCGACAGCTGGA 295
Qy      81 AspIleLeuSerThrTyrcysValaspAsnAsnGlnGlyProGlyGluAspGlyAla 100
Db      296 GACATCATTAACACTTAT-----GGGTCTGTGCGCCAGCACACGAGGAGGAGGCTCT 349
Qy      101 GlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrValAlaAsgAsn 120
Db      350 GCCAGGCGCAGTCAGCAGCTGAGATGCAATACCTACCTACACAC-----GAGGAT 400
Qy      121 GlyGluProGluProThrProValValAsnGlyGluGluSerLysGlyAspPro 140
Db      401 GGGGACTGTGAGGAAACAACTGAAGAGGCTGGAAGAGAGAACCCCTGCTCTGAGAGAGCCA 460
Qy      141 AsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHisArgArgProGln 160
Db      461 CCACCTGTCAAA-----GAGCCGCTGAGCAATAAGAGAGCAAAA-----TTG 502
Qy      161 GluLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsn 180
Db      503 GAAAGAAAATCCTAAAGAGATTAGCAAGAGCAACCTGCTAATGCAAAATCTGAAC 562
Qy      181 ThrLeuSerThrProGluGluLysLeuAlaAlaLeuLysLysLysThrAlaGluLeuLeu 200

```

Db 563 AAGTTGCAACACCCGGAAGAAAGTTGATTTTATTCAGAAGTATGCTGAATTCGCTG 622
 Qy GluGluHisArgAsnSerGlnLysGlnMetLysLeuGlnLysLysGlnSerGlnLeu 220
 Db 623 GATGAACATCTGACTGAGCAAAAGAGTTAAAGCTCTCTCCAAAGAAACACAGGTACAAAT 682
 Qy 221 ValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLys 240
 Db 683 CAAAGAAAGAGGACCGAGTTCAAGGTTGAACACAGCAGAGTATCTCTGCTCGAAGCAAA 742
 Qy 241 LeuGluSerLeuLysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGluGlyVal 260
 Db 743 TTGAGAGTCTGTGCGGAGCTCAGACACACAAAGACTCTGAAGGAGAGGCGCTT 802
 Qy 261 GlnArgAlaArgGluGluGluGluLysArgLysGluValThrSerHisPheGlnValThr 280
 Db 803 CAGCGGCGACGTGAGGAGAGAGAAAGAGAAATCAAGCCATTTCCAGATCC 862
 Qy 281 LeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGln 300
 Db 863 CTCAGGACATCCAGGCGAGATCGACGACGAGTGTGAGAAATATGAAGCTCTCTCAG 922
 Qy 301 GluAsnMetGluLeuAlaGluArgLysLysLysLysLysLysLysLysLysLysLys 320
 Db 923 GAGAACACAGAGCTTTGAGAAAGCTGAAAGAGATCATCATGATGATGATGATGATGATG 982
 Qy 321 GluHisIleAspLysValPheLysHisLysLysLysLysLysLysLysLysLysLys 340
 Db 983 GAGCATCTGGACAAATATTTAAACACAGAGACTGACGAGAGTGTGAGTGTGAGTGTG 1042
 Qy 341 LeuGlnGlnAlaGlnGlnMetLysLysLysLysLysLysLysLysLysLysLysLys 360
 Db 1043 CTTGAGAGGCGCCCAAGAAATGATGAAGAGAGCGGAGGAGCGACACACAGCAAGAAAG 1102
 Qy 361 PheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlu 380
 Db 1103 TATTGCTGAACGCGGACGAGAGTGTGAAACTTCAGCGGAAATGCTGAAGGAGCAAG 1162
 Qy 381 ThrHisLysLysGlnGlnLeuAlaLeuThrThrGluLysPheGluGluPheGlnAsnThr 400
 Db 1163 ACATCTGAGGCTGAGCTGACTCTCTACTAGGAGGTTTGAAGAAATTCAGAGACACA 1222
 Qy 401 LeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLys 420
 Db 1223 CTAACATAAGCAACGAGGTGTTGCCAGCTTCACACAGGAATGACACAAACACTAAG 1282
 Qy 421 LysIleLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 440
 Db 1283 AAAATGAAGAGCTGGAAGAGGACAGCCACACACACACACACACACACACACACACAC 1342
 Qy 441 LysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysLysLysLysLysLys 460
 Db 1343 AAAGCTCTGTGACATGATTGAAGAGAAAGCACTGAGAGCTAAAGATATGAGTCTTT 1402
 Qy 461 GlnValLysIleGlnArgLeuGluLysLysLysLysLysLysLysLysLysLysLys 480
 Db 1403 GTGATGAATATCGGAGGCTAGAGAACCTCTGCGGCTTTTACAGAGAGAGAAACGAA 1462
 Qy 481 LeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSerGly 500
 Db 1463 CTCACAAAAAATCAGAGACGCGAGAAATATCTGAAAGAGATGACCAAGTCAGCACAC 1522
 Qy 501 ProGluArgArgProGluGly 507
 Db 1523 TCCGATGAGAGCGCAGAGTCAAAAGCTCTCTGTGATCAAGAGATTGACGAGAGGAGTT 1582
 Qy 507 507
 Db 1583 AATAGTGTCCAAACCGCGTGAAGAAATCTGGCCACAGCCCTTCATGATTAATTCATCCA 1642
 Qy 507 507
 Db 1643 GAGTCAACCCCGCACCGAGTCCAAAGAAACCCAAACCGGAAATAGGCAGTTCTCAGGAGAGT 1702

Qy 508 -----ProGlyAlaGlnAlaProSerSerPro-----ArgValThrGlu 520
 Db 1703 GCTGACGCCCTCTCAAGAGCCAGACCAACCCCTCTGATCCCTTCACGGGATTCAGAG 1762
 Qy 521 AlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGlu 540
 Db 1763 AGTCCCTGCTCCCTAACTCTCAGGCTGAAGCCGAAAGGAGGAGGATGATCTGAACT 1822
 Qy 541 ProThrSerAla 544
 Db 1823 CCCTCCCAAGGCC 1834
 RESULT 26
 AB085905
 LOCUS
 DEFINITION Homo sapiens hMDP77 mRNA for muscle-derived protein 77, complete cds.
 ACCESSION AB085905
 VERSION AB085905.1 GI:35210317
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 REFERENCE
 1 Uyeda, A., Fujimori, K. and Taguchi, T.
 Molecular cloning of human homologue for MDP77
 Unpublished
 2 (bases 1 to 2869)
 Uyeda, A., Fujimori, K. and Taguchi, T.
 Direct Submission
 Submitted (29-MAY-2002) Atsuko Uyeda, AIST, Div. for Human Life
 Tech.; 1-8-31, Midorigaoka, Ikeda, Osaka 563-8577, Japan
 (E-mail:uyeda.atsuko@kaist.go.jp, Tel:81-727-51-9524,
 Fax:81-727-51-9628)
 FEATURES
 Location/Qualifiers
 1..2869
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="skeletal muscle"
 1..2869
 /gene="hMDP77"
 186..2220
 /gene="hMDP77"
 /codon_start=1
 /product="muscle-derived protein 77"
 /protein_id="BAC92688.1"
 /db_xref="GI:35210318"
 /translation="MEANHSQLSAEROSTPFGDSSSLPSHNGLEKEDGDSPTTPVP
 PKESVHPDISLEINROLEDINTYGAASATAGKSGASARASQEPENASPDNEDGDC
 ESTTEACRPEVAGCEPTVKEPVSNKEKLEKLLKLGKEANLLMKNLKLQTPPE
 KEDFLFKYVALLDHRTEQKLLKLOKQVIOKEKDLQGEHSRAILARSKLESLC
 RLQRHNTKLEALQRAEERKKRITSHFSQTLTDIQQIEQSERNNKLCQENT
 ELAKKSIIDQYELREHLDKIFKRELQKLVADKLEAQENMKAEERHREKEY
 LNKQAELKQAKVLKEQTLQALTLYSGRFEQFTLTKSNEVATFKQEMDKTT
 KMKLEKDTATKARFENCNKALIDMTIEKALRAKEVECFVMKIGLENLCALQEE
 RNELKTRDAIESEKDDQSNDSBPESNVSDQIDAEVNSVOTAVNTLAPM
 LTHPESTPHOSKETOPISGSSSADAALKEPPPLIPSRDSSEPLPLTPQABAE
 GSDAEPPSKASNSPAGTGAETQCEGLPVGQAQDQAKMKPEAPAGQAPTEASIQ
 KMSVDPAPACAAEHVAAYVFACEPSRQPPRAAABELPVGASGAPQPRNVATNLBG
 VD"

ORIGIN

Alignment Scores:
 Pred. No.: 7,8e-52 Length: 2869
 Score: 1275.00 Matches: 277
 Percent Similarity: 61.42% Conservative: 94
 Best Local Similarity: 45.88% Mismatches: 141
 Query Match: 45.78% Indels: 92
 DB: Gaps: 10

US-10-023-523-44 (1-546) x AB085905 (1-2869)

CS-10-023-523-44 (1-546) X AB5659505 (1-26997)

Qy	3	AsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerSerProGlyGln	22
Db	175	AATCATCTCTGAACAGCTCTCAGCGGAAACGACAGTCAACACCT	225
Qy	23	ProGluAlaGlyPro-----GluGlyAlaGlnGluArgProSerGlnAlaAlaProAla	40
Db	226	AGTTTCATCATCCAGTCAACATGGCTGGAGAGGAAGATGCCAG	273
Qy	41	ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla	60
Db	274	-----GATTCTCCAACCCAGTCCCAACCAACCAGAGAAA	306
Qy	61	ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu	80
Db	307	-----GAGGCAAGTGTGCAACCCGATATCTCTGAGAGCTGATTCACAGCTGGAA	357
Qy	81	AspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyVala	100
Db	358	GACATCATTAACACTTAT-----GGGTCTGCTGCCAGCAGCAGCAGGGAAGAGGGCTCT	411
Qy	101	GlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsn	120
Db	412	GCCAGGCCAGTCAGCAGCTGAGATGTCAGATCACTCAGCAAC	462
Qy	121	GlyGluProGluProThrProValValAsnGlnGlyLysGluProSerLysGlyAspPro	140
Db	463	GGGGACTGTGAGGAAACAACTGAAGAGCGTGGAGAGAACCCGTTGCTTCTGGAGAGCCA	522
Qy	141	AsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHisArgArgProGln	160
Db	523	CCCACTGTCTCAA-----GAGCCGTGACCAATGAGGAGCAAAA	564
Qy	161	GluLysLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsn	180
Db	565	GAAGAAGAAATCCTAAAGGATTAGGCAAGAAGCAACCTGCTAATGCAAAATCTCAAC	624
Qy	181	ThrLeuSerThrProGluGluLysLeuAlaAlaLeuCysLysLysTyrAlaGluLeu	200
Db	625	AAGTTGCAACACCGGAAGAAAGTTTGATTTTTTATTCAGAGATGATGCTGAATTCGTG	684
Qy	201	GluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeu	220
Db	685	GATGAACATCGTACTGAGCAAAAAGAGTTAAAGCTCTCCAAAAGAACAGGTACAAATT	744
Qy	221	ValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLys	240
Db	745	CAAAAAGAAAAGACCAAGTTACAGGTGGAACACAGCAGAGCTATCTCGCTCGAAGCAAA	804
Qy	241	LeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGluGlyVal	260
Db	805	TTGAGAGTCTGTGCCGGAGCTGCAGAGACACAACAAGACTCTGGAAGGAGGCGCTT	864
Qy	261	GlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnValThr	280
Db	865	CACGGGCACGTGAGGAGAAGAAAAGGAAGAAATCAAGCCCATTTCCAGAGTACC	924
Qy	281	LeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGln	300
Db	925	CTCAGGACATCCAGGGCCAGATCGAGCAGCAGAGTGGCGAAATATGAAGCTCTGTGAG	984
Qy	301	GluAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTyrGluLeuArgGlu	320
Db	985	GAGAAACACAGAGCTTGACAGAAAAGCTGAAAGCATCATCGATCATGATGAGCTCAGAG	1044
Qy	321	GluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeuValAspAlaLys	340
Db	1045	GAGCATCTGCACAAANAATTTAAACACAGAGAACTGCAGCAGAGCTGGTGGATGCAAG	1104
Qy	341	LeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluGluArgHisGlnArgGluLysAsp	360
Db	1105	CTTGAGAGGCCCAAGAAATCATGACGAAGACGGAGGAGGCACAAACAGAGAAAGGAA	1164

	Qy	361	PheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlu	380
	Dy	1165	TATTTGCTGAACAGGCAGCAGAGTGGAAATCTTCAGCGCAAGATGCTGTGAAGGAGCAAG	1224
	Qy	381	ThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThr	400
	Dy	1225	ACAGTCTGCGAGCTCAGCTCACTCTACTCAGGAAGGTTTGAGAATAATCCAGAGCAC	1284
	Qy	401	LeuSerLysSerSerGluAlaValPheThrThrPhoLysGlnGluMetGluLysMetThrLys	420
	Dy	1285	CTAACTAAAACAGCAACGAGGTGTTTGCCACGTTCAAACAGGAATATGGACAACAACTAAG	1344
	Qy	421	LysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArgTrpGluSerSerAsn	440
	Dy	1345	AAATGAAGAAGCTGGAAAGGACACAGGCCACATGGAAAGCCCATTGTGAGAACTGTAAAC	1404
	Qy	441	LysAlaLeuLeuGluMetAlaGluGlyLysThrValArgAspLysGluLeuGluGlyLeu	460
	Dy	1405	AAAGCTCTGTTGACATGATTTGAAGAGAAAGCACTGAGAGCTAAAGAATAATGATGCTTT	1464
	Qy	461	GlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluAla-gAsnAsp	480
	Dy	1465	GTGATGAANAATCGGAGGCTAGAGAACCTTCGCGTCTTTACAAGAAGAGAGAAACGAA	1524
	Qy	481	LeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSerGly	500
	Dy	1525	CTCCACAAAAAATCAAGACGCGAGAATAATCTGAAAAAGGATGCCAAAGTCAGCACAAAC	1584
	Qy	501	ProGluArgArgProGluGly-----	507
	Dy	1585	TCCGATCAAGACCGAGNATCAACCGTCTCTGTGGATCAAGAGATTGACGCGAGGAGGTT	1644
	Qy	507	-----	507
	Dy	1645	AATAGTGTCCAAACCGCGTGAAAAATCTGGCCACAGCCTTCATGATTAATCATCAPCCA	1704
	Qy	507	-----	507
	Dy	1705	GAGTCAACCCGACCAGTCCAAAGAAACCAACCCGAANTAGCGATTCTCAGGAGAGT	1764
	Qy	508	-----ProGlyAlaGlnAlaProSerSerPro-----ArgValThrGlu	520
	Dy	1765	GCTGACCGCCGCTCTCAAGAGCCAGAGCAACCCGCTTCGATCCCTTCACCGGATTCAGAG	1824
	Qy	521	AlaProCystYrProGlyAlaProSerThrGluAlaserGluInThrGlyProGlnGlu	540
	Dy	1825	AGTCCCTCGCTCCCTTAATCTCTCAGGCTGAAGCCGAAGGCGAGTGATGCTGAACCT	1884
	Qy	541	ProThr-SerAla 544	
	Dy	1885	CCCTCCAAGGCC 1896	

RESULT 27

HSM805276
YOGIA

LOCUS
DEFINITION

5
4
4
4
4
4
4
4
4
4

ACCESSION

VERSION

KEYWORDS
SOURCE

ORGANISM

Figure 1

REFERENCE

DISCUSSION

TITLE

HSM805276 4631 bp mRNA linear PRI 17-JUN-2003
 Homo sapiens mRNA; cDNA DKFp451F022 (from clone DKFp451F022);
 complete cds.
 AL834248 AL834248
 AL834248.2 GI:30268572

 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 4631)
 Wambutt,R.K., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
 Robo,G., Han,M. and Wiemann,S.
 Direct Submission
 Submitted (17-JUN-2003) MIPS, Ingolstaedter Landstr.1, D-85764
 Neuherberg, GERMANY
 On Apr 30, 2003 this sequence version replaced gi:21739792.
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
 consortium of the German Genome Project.
 This clone (DKFZp451F022) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
 information about the clone and the sequencing project is available
 at http://mips.gsf.de/proj/cDNA/.

FEATURES

source

Location/Qualifiers
 1..4631
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="RZPD:DKFZp451F022"
 /db_xref="taxon:9606"
 /map="6Q24.1"
 /clone="DKFZp451F022"
 /tissue_type="human skeletal muscle"
 /clone_lib="451 (synonym: hicc1). Vector pSport1; host
 DH10B; sites NotI + SalI"
 /dev_stage="adult"
 1..4631
 /gene="DKFZp451F022"

gene

CDS

1..2148
 /note="similar to muscle-derived protein MDP77 variant
 1 (Mus musculus), differentially spliced"
 /codon_start=1
 /product="hypotheetical protein"
 /protein_id="CAD38924.2"
 /db_xref="GI:30268573"
 /db_xref="SPTRMBL:Q8N3L3"
 /translation="GCCSSDLPPASFLSLFLPLSPDTTPKMEANHSEQLSABR
 OSTPGDSSLPNSHNGLEKEDGOSPTVQPEKEASVHPDISELNARQLEDINTYGV
 SAASTAGEGSARASQEPENAESPDNEDGCEETEEZAGRPVAGSPPTVKFVSNK
 EQLKLLGLGKGNLQNLQNPTEKFDPLFKYAELODERTKQKLLQ
 KQVQIQEKQLOQEHRAILARSLKSLRELOHNKTLKEALQARAREEKKRKE
 ITSPQSTLTIQGGIEQSRNKLQENTELAEKLKIIIDQVLEEHLDKIFKPR
 ELQQLVDALKEAQEWKAEERHREKEYLQNAEWKLQAKVLEQETVLQAOIT
 LVSGRFPOSTLTSNVEVFTFKQEMDKTKKMKLEKDTATWKAIFENCNKALLDM
 TEELKRALKEVCFWKIYLENLCLALQEBNRLHKIRDAEISEKDDOSQHNDEE
 PESNVSVQEDAEVNSVOTAVKNLATAPMLIHPESTPHQSKETQPEIGSSQESAD
 AALKEPFPPIIPSRDSESPUPPTQAEAGGSDAEPFSPKSNPSRAGLAEBOCEGL
 PVGAQDAQAEKPEARSQAPQAPTEASLQKMEADVPAPACAAEEHVAAVPAPEPS
 RQPPFAAAEELPVGASAGPQPRNVADTNLEQVD"
 4589..4594
 /polya_signal
 4612
 /polya_site
 /gene="DKFZp451F022"

ORIGIN

Alignment Scores:

Pred. No.: 1.28e-51 Length: 4631
 Score: 1275.00 Matches: 277
 Percent Similarity: 61.42% Conservative: 94
 Best Local Similarity: 45.88% Mismatches: 141
 Query Match: 45.78% Indels: 92
 DB: 9 Gaps: 10

US-10-023-523-44 (1-546) x HSM805276 (1-4631)

Qy 3 ArgGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerProGlyGln 22
 Db 103 AATCACTTGAACAGCTCTCAGCGGAACGACAGTCAACACTT-----CCAGGTGAC 153
 Qy 23 ProGluAlaGlyPro-----GluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
 Db 154 AGTTTCATCATCCAGTCACCAATGGCTGGAGGAGGAAGATGGCCAG----- 201
 Qy 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
 Db 202 -----GATTCTCCAAACCCAGTCACCAACCCAGAGAAA----- 234
 Qy 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80

Db 235 -----GAGCAAGTGTGCACCCGATATCTCTGAAGAGCTGAATCGACAGCTGGAA 285
 Qy 81 AspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAla 100
 Db 286 GACATCATTAACACTTAT-----GGGTCTGTGCCAGCACAGCAGGGAAGAGGGCTCT 339
 Qy 101 GlnGlyGluProLagLupProGluAspAlaGluLysSerArgThrTyrValAlaArgAsn 120
 Db 340 GCCAGGGCCAGTCAGCGCTGAGAAATGAGAAATCAGAAATCACCCTGACAAC-----GAGGAT 390
 Qy 121 GlyGluProGluProThrProValValAsnGlyGluLysGluProSerLysGlyAspPro 140
 Db 391 GGGGACTGTGAGCAACAACCTGAGAGGCTGGAGAGAACCCCGTTCCTTCGAGAGGCCA 450
 Qy 141 AsnThrGluGluLeuArgLnsSerAspGluValGlyAspArgAspHisArgProGln 160
 Db 451 CCCACTGTCAAA-----GAGCCCGTCAGCAATAAGAGGCAAAAA-----TTG 492
 Qy 161 GluLysLysLysAlaLysGlyLeuGlyLysGluLeuThrLeuLeuMetGlnThrLeuAsn 180
 Db 493 GAAAGAAATCTTAAAGATTAGCAAGAAAGCCCACTGCTATGCAAAATCTGAAC 552
 Qy 181 ThrLeuSerThrProGluLysLeuAlaLeuLysLysLysTyrAlaGluLeuLeu 200
 Db 553 AAGTTGCAAAACACCGGAGAAAAGTTGATTTTATTCAAGAAGTATGCTCAATGCTG 612
 Qy 201 GluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuLysLysGlnSerGlnLeu 220
 Db 613 GATGAATCATGCTACTCAGCAAAAGAAAGTTAAAGTCTCTCCAAAGAAACAGGTACAAT 672
 Qy 221 ValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLys 240
 Db 673 CAAAAGAAAAGACCCAGITACAGGTGAACAGCAGAGACTATCTCGCTCGAAGCAA 732
 Qy 241 LeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGlyVal 260
 Db 733 TTGGAGAGTCTGTGCGGGAGCTGCGAGAGACACAAAGACTCTGAAGGAAGAGCGCTT 792
 Qy 261 GlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnValThr 280
 Db 793 CACGGGACGCTGAGAGAGAGAGAAAAGAGAGAAATCACAAGCCATTTCAGAGTACC 852
 Qy 281 LeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGln 300
 Db 853 CTCACGGACATCCAGGGCCAGATCGAGCAGCAGAGTGAAGAAATATGAAGCTCTGTCTAG 912
 Qy 301 GluAsnMetLeuAlaGluArgLysLysLysLeuLeuGluGlnTyrGluLeuArgGlu 320
 Db 913 GAGAACACAGAGCTTCGAGAAAGCTGAAAAGCATCATCGATGATGAGCTCAGAGAG 972
 Qy 321 GluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeuValAspAlaLys 340
 Db 973 GAGCATCTGACAAAAATATTTAAACACAGAGAACTGCAGCAGAGCTGGTGGATGCAAG 1032
 Qy 341 LeuGlnGluAlaGlnLeuMetLysGluAlaGluGluArgHisGlnArgGluLysAsp 360
 Db 1033 CTTGAGCAGGCCCAAGAAATGATGAAGGAAGCGGAGGAGCAGACACAAACGAAAGGAA 1092
 Qy 361 PheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGln 380
 Db 1093 TATTTGCTGAACCCAGGCAGCAGAGTGGAAATTCAGCGGAAAGTGTGAGGAGCAAGAG 1152
 Qy 381 ThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThr 400
 Db 1153 ACAGTCTCTCAGCTCAGCTCACTCTCTACTCAGGAAGTTTGAAGAAATTCAGAGCACA 1212
 Qy 401 LeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLys 420
 Db 1213 CTAACCTAAAGAACACGAGGTGTTGCCAGCTTCAACAGAGAAATGGACAAACACTAAG 1272
 Qy 421 LysIleLysLysLeuGluLysGluThrMetTyrArgSerArgThrGluSerSerAsn 440

```

Db 1273 ABAATCAAGAGCTGGAAGAAAGGACACAGCCACATGTAAGAAACCCGATTGAGAACTGTGAAC 1332
QY 441 LysAlaLeuGluMetAlaGluGlyThrValArgAspLysGluLeuGluGlyLeu 460
Db 1333 AAAGCTCTCTTGGACATGATTGAGAGAAAGCACTGAGAGCTAAAGATATGAGTGCITT 1392
QY 461 GlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAsp 480
Db 1393 GTGATGAAATCGGAGGCTAGAGAACCTCTCCGCTGCTTTACAAGAGAGAGAAACGAA 1452
QY 481 LeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSerGly 500
Db 1453 CTCACAAAATATCAGACAGCCAGCAATATCTGAAAGGATGACCAAGTCAGCACAAAC 1512
QY 501 ProGluArgProGluGly 507
Db 1513 TCCGATGAGAGCCAGAGTCAAAACCTCTCTGTGGATCAAGAGATTGACCGAGAGAGGTT 1572
QY 507 507
Db 1573 ATAGTGTCCTCAACCGCTGTAATAATCTGGCCACAGCCCTTCATGATTAATTCATCATCA 1632
QY 507 507
Db 1633 GAGTCAACCCCGACAGTCCAAAGAAACCCAAACCGAAATAGGAGTCTCTCAGGAGAGT 1692
QY 508 -----ProGlyAlaGlnAlaProSerSerPro-----ArgValThrGlu 520
Db 1693 GCTGACGCGCTCTCAAGGAGCCAGGACCAACCCCTCTGATCCCTTCACGGGATTCAGAG 1752
QY 521 AlaProCysTyProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGlu 540
Db 1753 AGTCCCTCCCTCCCTCACTCTCAGGCTGAGCGGAGGAGGAGTGTGCTGAACT 1812
QY 541 ProThrSerAla 544
Db 1813 CCTCCAAAGGCC 1824

RESULT 28
AC110537/c
LOCUS
DEFINITION AC110537 242048 bp DNA linear HTG 24-AUG-2002
pieces.
AC110537
VERSION AC110537.5 GI:22475046
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 242048)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Mus musculus, clone RP23-44A4
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 242048)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepe,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Gilde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagob,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRoque,K., Lamazares,R.,
Lander,S., Lehoczy,J., Levine,R., Liu,G., MacLean,C.,
MacDonald,T.P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Norbu,C.,
Nusbaum,C., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Rietman,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,

```

```

Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (13-FEB-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 242048)
Birren,B., Bustien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepe,Y., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagob,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., MacLean,C., MacDonald,T.P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mienga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 24, 2002 this sequence version replaced gi:20531889.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L22330
Center clone name: 44 A 4
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 237039 bases at least Q40
Consensus quality: 239586 bases at least Q30
Consensus quality: 240394 bases at least Q20
Insert size: 210000; agarose-fp
Insert size: 240848; sum-of-contigs
Quality coverage: 12.5 in Q20 bases; agarose-fp
Quality coverage: 10.9 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1 28688 28787: contig of 100 bp
* 28788 30488: contig of 1701 bp in length
* 30489 30588: gap of 100 bp
* 30589 110109: contig of 79521 bp in length
* 110110 112009: gap of 100 bp
* 110210 113706: contig of 3497 bp in length
* 113707 113806: gap of 100 bp
* 113807 120823: contig of 7017 bp in length
* 120824 120923: gap of 100 bp
* 120924 133081: contig of 12158 bp in length

```

																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																											</
--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

Db	87844	AAAGCATATCTCATGATGATGCTAGAGGACATCAAGAAGGACTTTTAATCACTCACTTAAA 87785
QY	57	GlyAlaGlnAlaArgThrAlaGlnSer---GlyAlaLeuArgAspValSerGluGluLeu 75
Db	87784	GAATATCAGAGAGACACTGCTAAAGAGGTACAGTCCTTAAGAAAAACAGGAAACACA 87725
QY	76	SerArgGlnLeuGluAspIleLeuSerThrTyrcysValaspAsnAsnGlnGlyGlyPro 95
Db	87724	ATCAAAACAGGTAGAA-----GTCCTTACAGAAAAAGAGAAAAACATACAAACTGGTGAT 87671
QY	96	GlyGluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThr 115
Db	87670	GGAAATGAACAAACCACTAGACCTAAAGGGAGAGTAGACACATAAAGAAACTCA 87611
QY	116	TyrValalaArgAsnGlyGluProGluProThrProValVal----- 129
Db	87610	AAGTGAGGCAACACTGAGATAGAAACCCCTAGAAAGAAATCTGGAACCATAGATTGAG 87551
QY	130	-----AsnGlyGluLysGluProSerLysGlyAsp----- 139
Db	87550	CATCAGCACAGAATACAGAGATGGAGGTACTTGAGGAGGCTAGTGAGGGCGGACGGG 87491
QY	140	-----ProAsnThrGlu----- 143
Db	87490	CGACGGGTCCCCCTTTCCTTCAGAAAGTCAACGATAGTTGGGGAGCAGGTGGGGGT 87431
QY	143	----- 143
Db	87430	GGTTGTGACTGAAATGGATCTGTTGAAATTTGGCAATGGAAGAGCTAGAAATTTGTGG 87371
QY	144	-----GluIleArgGlnSerAspGluValGlyAspArgAspHisArg 158
Db	87370	GTTAGGAGTAAAGACACACATGAGAGAAAGAGAAAGAAATTCCTGGACGAGAGCTGAC 87311
QY	158	g---ProGlnGlu-----LysLysLysAlaLysGlyLeuG 169
Db	87310	AGTCTCTCTGATGGCCAGCAGATTACAGATGTCAGCAGAGCAAAAGAGAGACCTTAG 87251
QY	169	YlysGluIleThrIleuLeuMetGlnThrIleuAsnThrLeuSerThrProGluGluLysLe 189
Db	87250	AAAAGAAGTTTGTGATGATCAAGCGCTAAACACCCCTTCAACCCAGAGAGAGAGCT 87191
QY	189	uAlaAlaLeuCysLysLysTyrrAlaGluLeuGluGluHisArgAsnSerGlnLysG 209
Db	87190	GGCAGCTCTCTGTAAAGAAATATGCTGATCTCTCTGGAAGAGAGCAGGAATGTTCA 87131
QY	209	nMetLysLeuGlnLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgG 229
Db	87130	AATGAAGATTCTGCAGAAAGCAAGCCAGATTTGTGAAGAGAAAGTTCCACCTTCAGAG 87071
QY	229	YGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLysLeuCysArgGluLeuG 249
Db	87070	TGAACACAGCAAGCCATCTTGGCAGAGCAAACTGGATCTCTTTGCGAGGAACCTTCA 87011
QY	249	nArgHisAsnArgSerLeuLysGluGlyValGlnArgAlaArgGluGluGlnGly 269
Db	87010	CGGTCAATAAAGACCTTAAAGAGAGAGAAACATGCAGCAGCAGCAGAGAGAGAGAACG 86951
QY	269	sArgLysGluValThrSerHisGlnValThrLeuAsnAspIleGlnLeuGlnMetG 289
Db	86950	ACGGAAGAGAGCAACAGCATTTCCAGATACTCTTAATGAATCCACAGCGCAGTTGGA 86891
QY	289	uGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLe 309
Db	86890	ACAACATGACATCCACAATGCCAACTGCACAGGAGAAACATTGAATGGGAGAGAGATT 86831
QY	309	uLysLysLeuIleGluGlnThrGluLeuArgGluGluHisIleAspLysValPheLysH 329
Db	86830	GAAGAGCTTATTGAGCAGATGTCATTAAGGAGAGAGATATTGATTAAGTATTCAACA 86771
QY	329	sLysAspLeuGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnLeuMetLeuL 349
Db	86770	CAAGGAATTGCAACAACAGCTTTGGATGCCAACTTCAGCAACCAACACAGCTGATAAA 86711


```

Db      813  CAGCGCGCACGTGAGGAAGAGAGAAAGGAATATCAAGCCATTTCCAGAGTACC 872
QY      281  LeuAenAspIleGlnLeuGlnMetGluGlnHisGlnGluArgAsnSerLysLeuArgGln 300
Db      873  CTCAGGACATCTCAGGGCCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 932
QY      301  GluAenMetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTyrGluLeuArgGlu 320
Db      933  GAGACACAGAGCTTGCAGAAAGCTGAAAGCACCATCGATCATGATGAGCTCAGAGAG 992
QY      321  GluHisIleAspLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLys 340
Db      993  GAGCATCTGACAAAATATTAACACAGAGAACTGCAGAGAACTGCTGTGATGCAAG 1052
QY      341  LeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluGluArgHisGlnArgGluLysAsp 360
Db      1053  CTTGAGCAGGCCCAAGAAATGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 1112
QY      361  PheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlu 380
Db      1113  TATTTGCTGAACGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1172
QY      381  ThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThr 400
Db      1173  ACAGTCCTCAGAGCTCAGCTACTCTCTACTCAGGAAGGTTTGAAGAATTCAGAGACA 1232
QY      401  LeuSerLysSerSerGluValPheThrThrPheLysGlnGlnMetGluLysMetThrLys 420
Db      1233  CTAACCTAAGGCAACGAGGTGTTGCCAGCTTCAACAGAGAAATGCACAAACAACTAAG 1292
QY      421  LysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArgTyrGluSerSerAsn 440
Db      1293  AAAATGAAGAGCTGGAAGAGACACAGCCACCAAGGAGGAGGAGGAGGAGGAGGAGG 1352
QY      441  LysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGluLeuGluGluLeu 460
Db      1353  AAGGCTCTGTGACATGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1412
QY      461  GlnValLysIleGlnArgLeuGluLysLysCysArgAlaLeuGlnThrGluArgAsnAsp 480
Db      1413  GTGATGAAATCTCGGAGGCTAGAGAACTCTCCGCTGCTTTACAGAGAGAGAGAGAGAA 1472
QY      481  LeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSerGly 500
Db      1473  CTCACAAATAATCAGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1532
QY      501  ProGluArgArgProGluGly 507
Db      1533  TCCGATGAAGAGCCAGAGTCAACCGTCTCTGTGGATCAAGAGATTGACGAGAGAGGTT 1592
QY      507  ----- 507
Db      1593  AATAGTGTCAAAACCGCGTGAAATAATCTGGCCACAGCTTCATGATTAATCATCATCA 1652
QY      507  ----- 507
Db      1653  GAGTCAACCCGACAGTCCAAAGAAACCCAGAAATAGGAGGAGTCTCAGGAGAGT 1712
QY      508  -----ProGlyAlaGlnAlaProSerSerPro-----ArgValThrGlu 520
Db      1713  GGTGACGCGCTCTCAAGGAGCCAGAGCAACCCCTCTGATCCCTTCACGGGATTCAGAG 1772
QY      521  AlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGlu 540
Db      1773  AGTCCCTCGCTCCCTAACTCTCTAGGCTGAGCGGAGGAGGAGGAGGAGGAGGAGGAG 1832
QY      541  ProThrSerAla 544
Db      1833  CCTCCAGAGCC 1844

```

RESULT 30
D89999

```

LOCUS      D89999          4704 bp      mRNA      linear      VRT 26-APR-2000
DEFINITION  Gallus gallus MDP77 mRNA for muscle derived protein, complete cds.
ACCESSION   D89999
VERSION     D89999.1  GI:7619883
KEYWORDS    Gallus gallus (chicken)
SOURCE      Gallus gallus
ORGANISM    Gallus gallus
REFERENCE   1  (sites)
AUTHORS     Ueda, A., Fukui, I., Fujimori, K., Kiyosue, K., Nishimune, H., Kasai, M.
and Taguchi, T.
TITLE       MDP77: A novel neurite-outgrowth-promoting protein predominantly
expressed in chick muscles
JOURNAL     Biochem. Biophys. Res. Commun. 269 (2), 564-569 (2000)
MEDLINE     20175243
PUBMED      10708594
REFERENCE   2  (bases 1 to 4704)
AUTHORS     Ueda, A.
DIRECT SUBMISSION
TITLE       Submitted (24-DEC-1996) Atsuko Ueda, Osaka National Research
Institute, Department of Organic Materials; 1-8-31, Midorigaoka,
Ikeda, Osaka 563-8577, Japan (E-mail: au-ueda@onri.go.jp,
Tel: +81-727-51-9524, Fax: +81-727-51-9628)
JOURNAL
FEATURES    Location/Qualifiers
             source
             1..4704
             /organism="Gallus gallus"
             /mol_type="mRNA"
             /db_xref="taxon:9031"
             /sex="male"
             /tissue_type="muscle"
             /dev_stage="postnatal 8 days"
             1..4704
             /gene="MDP77"
             /genes="MDP77"
             /function="neurite-outgrowth-promoting protein"
             /notes="old name: muscle derived factor F460
             with molecular mass 77 k"
             /codon_start=1
             /product="muscle derived protein"
             /protein_id="BAA94755.1"
             /db_xref="GI:7619884"
             /translation="MENDQFTKQQVTTPTQDNQGSKAEFVPSQPLSPNTQSA
             OPMACTDISELNROLEDIKTYGSAASLVEKGTATETDKPEKEDVGSMDACED
             VNEESKDKPACDASRAKESPSAEKQLEKILKLGKEATLLMQSLNKLTPSEKL
             DLLFKYALLESRAEQQLYLOKQAQITKENDQLOSEHSRAILARSKLSCRE
             LQHNKTLKEETIQRAREDEKKEITNHFQGTLSLSEIQAIQEQSERNNKLCQNTL
             AEKLSIIDYELREHLDKIFRRELQQLVDKLEQSQEMKRAERHQEKEXYLL
             NQAAEWKLAQMLKEQETVLAQITLYSERFEFQKTLTKSNEVFATFKQEMKTK
             MKKLEKDTATWKSFRNCNRLIMIEEKAMTKYECFVLKIQLENLICRALQBERN
             ELVRKIKOALPEEVNGNDILREDDANTNPSSQASIELCAADKNMLQELAEAFV
             SHKAEETLPSDGSNPETCNVQCEATISVPELPSHLTSPQEAHNCQFQSMSTSAPEH
             MPATENMTLIENMEKPSMPPEMVPETPESVPIPTGVPTPKIMPATPESVP
             TLMQNTSAPLGNMNPASTKPKRAVEHVDIIEFLIPDQPAEQKQGTDMNAVD"
ORIGIN
Alignment Scores:
Pred. No.:      1,78e-50      Length:      4704
Score:          1251.00      Matches:    276
Percent Similarity: 65.52%      Conservative: 85
Best Local Similarity: 50.09%      Mismatches: 135
Query Match:    44.92%      Indels:    55
DB:              5              Gaps:      12

US-10-023-523-44 (1-546) x D89999 (1-4704)
QY      12  LysGlnSerAsnProLysSerSerSerGlnProGluAlaGlyProGluGlyAlaGln 31
Db      220  AAGCAGCAGCAGGTTACCTACCTCCTACCTCAGCAGACACACAGGGG----- 264
QY      32  GluArgProSerGlnAlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAla 51

```


Db 265 CAAGCAAGCGGAACTGTCCTCCAGAGCCCTTTCCGCC---ACAAACCAACC 321
Qy 52 ProArgLysProGluGlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspVal 71
Db 322 AGTGCCAGCGCTGAGATGGCCACGTGT-----GACATC 354
Qy 72 SerGluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsn 91
Db 355 TCAGAGAGCTGAAACCGGCACTCGAGGACATATTAAACATAT----- 399
Qy 92 GlnGlyGlyProGlyGluAspGlyAla-----GlnGlyGluProAlaGluPro 107
Db 400 -----GGTCTGCGAGCGAGTCTGTGTAGAGAGGAAGCACTACAGCAAACT 447
Qy 108 GluAspAlaGluLysSerArgThrTyrValAlaArgAsnGlnGluProGluProThrPro 127
Db 448 GACAAAGCCAGAGGAAGAGATGTGGCGAGTGTGAGGATGAGGAT----- 501
Qy 128 ValValAsnGlyGluLysGluProSerLysGlyAspProAsnThrGluGluIleArgGln 147
Db 502 ---GTAATGAGAAAGTGAGAAAGACAAACAGCTCTGGA---GATGCTTCAAGAGCA 555
Qy 148 SerAspGluValGlyAspArgAspHisArgArgProGlnGluLysLysLysAlaLysGly 167
Db 556 AAGGAGCCAGTGCAGCAAGCAAAAG-----CTGGAGAGAAATCTCGAAAGGA 609
Qy 168 LeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGlu 187
Db 610 CTAGGGAAGAGAGCCCTCTCTATGCAAGCTTGAACAGCTGACTACTCTCCAGAGGAG 669
Qy 188 LysLeuAlaAlaLeuCysLysLysTyrAlaGluLeuLeuGluGluHisArgAsnSerGln 207
Db 670 AAGCTGGACCTGTTATTAAAGAGATGCTGAGTTGCTTGAGGAGCATCGTGCTGAGCAG 729
Qy 208 LysGlnMetLysLeuLeuLysLysGlnSerGlnLeuValGlnGluLysAspHisLeu 227
Db 730 AAGCAGCTCAAGTACTCTCAGAGAGAGAGCGCCAGATCCAGAGGAAGACCAAGT 789
Qy 228 ArgGlyGluHisLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGlu 247
Db 790 CAGAGTGAGCAGCGGAGCCATCTCTGTCAGCAAGCTTGAGAGCTGTGCGCGGAG 849
Qy 248 LeuGlnArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGluGlu 267
Db 850 CTCAGAGCAACAACAAACCTCAAGGAGGAACAATTCAGCGGCGCAGGAGGAAGAT 909
Qy 268 GluLysArgGlyGluValThrSerHisGlnValThrLeuAsnAspIleGlnLeuGln 287
Db 910 GAGAGAGGAAGAATAACAATCATTTCCAGGCGCAGCTGAGTGAATCCAGGCTCAG 969
Qy 288 MetGluGlnHisAsnGluArgAsnSerLysLeuA-gGlnGluAsnMetGluLeuAlaGlu 307
Db 970 ATTGAGCAGCAAAAGTGAGAGGAACAATGAGCTCTGCCAGGAGAACACAGAGCTGCGAG 1029
Qy 308 ArgLeuLysLysLeuIleGluGlnTyrGluLeuArgGluGluHisLysLysValPhe 327
Db 1030 AAGCTGAAAGCATTAATCAGCAGTATGAGTGCAGGGAAGACACCTTGCAAAATATT 1089
Qy 328 LysHisLysAspLeuGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMet 347
Db 1090 AAGCAGAGAACTTCAACAGAAATTTGTTGGATGCCAAGTTGGAGCAGTCTCAGGAATG 1149
Qy 348 LeuLysGluAlaGluArgHisGlnArgGluLysAspPheLeuLeuLysGluAlaVal 367
Db 1150 ATGAGGAAGCAGAGGAGGACATCAGAGGAAGAAAGGAATATCTCTGTAATCAAGCGCA 1209
Qy 368 GluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLeuLysGlnGlnLeu 387
Db 1210 GAATGAAGCTACAGGCCAAATGTTAAGGACAGACAGACAGTCTGCGACACAGATC 1269
Qy 388 AlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerGluVal 407

Db 1270 ACTCTCTACTCTGAGAGATTTGAAGAAATTCAGAAAAACATTGACCAAAAGCAATGAGTG 1329
Qy 408 PheThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLysLysLys 427
Db 1330 TTGTCTACCTTTCAACAGAGATGGGAAATGACAAAGAAATGAAGAAGTTGAAAAG 1389
Qy 428 GluThrThrMetTyrArgSerArgTrpGluSerSerAsnLysAlaLeuLeuGluMetAla 447
Db 1390 GATACTGTCTACATGGAAATCCAGGTTTGAGAAATGTAACAGAGCATTACTGGACATGATT 1449
Qy 448 GluGluLysThrValArgAspLysGluLeuGluGlyLeuGlnValLysLysLysLys 467
Db 1450 GAAGAGAAAGCCATGAGGACCCAGGAAATACGATGCTTGTGCTGAAATCCAGAGGCTA 1509
Qy 468 GluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLysLysLysLysLysLys 487
Db 1510 GAGAACCTTTGCGAGCTCTGAGGAGAGAGAGATGAATGTACAGAAAAATAAACAA 1569
Qy 488 Leu-----SerAlaGlyGlyGlnGlySerLeu---ThrAspSerGlyProGlu 502
Db 1570 GCACAGCTCCCTGAAGAGGTGAATGGAATAGCATCTTAGAAGAGCAGCAGTGCAT 1629
Qy 503 ArgArgProGluGly-ProGlyAlaGlnAlaProSerSerProArgValThrGluAlaPr 522
Db 1630 ACAACCCCTTCCTTCCGAGCAGGAGCAAGCATTTGAGCTATGCTGCTGCAAGAAACATG 1689
Qy 522 cOys-----TyrProGlyAlaProSerThrGluAlaSerGlu 534
Db 1690 CTGAGAGGAGTGTGAAAGCTTTTCAGGGTATCCC-----ACAAA 1728
Qy 534 yGlnThrGlyProGlnGluProThrSerAla 544
Db 1729 GCAGAGAGAGCCCTCCAGCAGCGGAGCA 1759

RESULT 31
LOCUS E26533 4723 bp DNA linear PAT 18-JUN-2001
DEFINITION Protein with neurite outgrowth activity.
ACCESSION E26533
VERSION E26533.1 GI:13026200
KEYWORDS JP 1999147897-A/1.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 4723)
AUTHORS Takahisa,T.
TITLE Protein with neurite outgrowth activity
JOURNAL Patent: JP 1999147897-A 1 02-JUN-1999;
AGENCY OF IND SCIENCE & TECHNOL
COMMENT OS Unidentified
PN JP 1999147897-A/1
PD 02-JUN-1999
PF 13-NOV-1997 JP 1997331242
PR TAKAHISA TAGUCHI
PI C07K14/52,A61K38/00,A61K38/00,A61K38/00,C12N15/09,A61K37/02,
PC A61K37/02,
PC A61K37/02,C12N15/00
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..4723 /organism='Unidentified'.
FEATURES
source Location/Qualifiers
1..4723 /organism='unidentified'
/mol_type='genomic DNA'
/db_xref='taxon:32644'

ORIGIN
Alignment Scores: 1.78e-50 Length: 4723
Pred. No.: 1251.00 Matches: 276
Score:

Percent Similarity:	65.52%	Conservative:	85
Best Local Similarity:	50.09%	Mismatches:	135
Query Match:	44.92%	Indels:	55
DB:	6	Gaps:	12

US-10-023-523-44 (1-546) x E26533 (1-4723)

QY	12	LysGlnSerAsnProLysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGln	31
DB	220	AAAGCAGCAGCAGCTTACACACTTCACTACGACAGACAACACAGGGG	264
QY	32	GluArgProSerGlnAlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAla	51
DB	265	CAAAGCAGGCGGAACCTGTCCTCGTCTCAGAGCCCTTTCCCCC--ACAAACCAAC	321
QY	52	ProArgLysProGluGlyValGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspVal	71
DB	322	AGTGCCAGCAGCTGAGATGCCACCTGT--GACATC	354
QY	72	SerGluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsn	91
DB	355	TCAGAGGAGCTGAACCGCAACTCGAGACATTATTAAACATAT--	399
QY	92	GlnGlyGlyProGlyGluAspGlyAla--GlnGlyGluProAlaGluPro	107
DB	400	-----GGGTCTGCAGCGAGTCTGTAGAGGAAGGCACTACAGCAGAACT	447
QY	108	GluAspAlaGluLysSerArgThrTyrValAlaArgAsnGlyGluProGluProThrPro	127
DB	448	GACAAGCCAGAGAAGGAAGATGTGGCAGTATGAGGATGACAGTGTGAGGAT--	501
QY	128	ValValAsnGlyGluLysGluProSerLysGlyAspProAsnThrGluGluIleArgGln	147
DB	502	---GTAAATGAAGAAGTGAGAAAGACAACCAGCTCTTGGA---CATGCTTCAAGACA	555
QY	148	SerAspGluValGlyAspArgAspHisArgArgProGlnLysLysLysAlaLysGly	167
DB	555	AAAGGACCCAGTCGCCAGCAAGAAACAAAG-----CTGGAGAAGAAATCTCTGAAAGGA	609
QY	168	LeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGlu	187
DB	610	CTAGGGAAGGAAGCACCCTCTCATCAAGCTTGGAACAGCTGACTACTCCAGGAG	669
QY	188	LysLeuAlaLeuCysLysLysTyrAlaGluLeuGluGluHisArgAsnSerGln	207
DB	670	AAGCTGGACCTGTTATTATAAGAGATGCTGAGTTGCTTGAGGAGCATCGTCTCGACG	729
QY	208	LysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeu	227
DB	730	AAGCAGCTCAAGTACCTGCGAAGAGAGCAGCGCCAGATCATCAGGAGGAAGACCAAGTTG	789
QY	228	ArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGlu	247
DB	790	CAGAGTGAGCACACCCGAGCCATCTTGCTCGCAGCAAGCTTGAGAGCCTGTGCGCGGAG	849
QY	248	LeuGlnArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGluGluGlu	267
DB	850	CTCAGAGGCCAACAAACCCCTCAAGGAGGAACAAATTCAAGCGGCGGAGGAAGAT	909
QY	268	GluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGln	287
DB	910	GAGAAGAGGAAGAAATAACAATCATTTCCAGGCGACGCTGAGTGAATCCAGGCTCAG	969
QY	288	MetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGluLeuAlaGlu	307
DB	970	ATTGAGCAGCAAGATGAGAGAGAAATCAAGACTCTGCGAGGAGAAACACAGAGCTGCGAG	1029
QY	308	ArgLeuLysLysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAspLysValPhe	327
DB	1030	AAGCTGAAGCATATTATGACCAAGTATGAGCTGCGGGAAGACACCTTGACAATAATATT	1089
QY	328	LysHisLysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnAlaGlnLeuMet	347

Db	1090	AAGCACAGAGAACTTCAACAGAAATTTGGTGGATGCCAAGTTGGAGCAGTCTCAGGAAATG	1144
Qy	348	LeuLysGluAlaGluGluA ^{ArgHis} GlnA ^{Arg} GlyLeuLysAspPheLeuLeuLysGlnAlaVal	367
Db	1150	ATGAAGGAGCAGAGGAGCGACATCAGAAGAAAAGGAATATCTCTGATCATCAGGCGCA	1209
Qy	368	GluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLeuLysGlnGln	387
Db	1210	GAATGGAAGTACAGGCCAAATGTTAAAGGAACAGGACAGACAGTCTCTGAGGCACAGATC	1269
Qy	388	AlaLeuThrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluVal	407
Db	1270	ACTCTTACTCTCAGAGATTGGAAGAAATTCAGAAACATTGGACCAAAAGCAATGAAGTG	1329
Qy	408	PheThrThrPheLysGlnGluMetGluLysMetThrLysLysLeuLysLysLeuGluLys	427
Db	1330	TTTGCTACCTTCAACAGGAGATGGAGAAAATGACAAAGAAAATGAAGTTGGAAGAAG	1389
Qy	428	GluThrThrMetTyrArgSerArgTrpGluSerSerAsnLysAlaLeuLeuGluMetAla	447
Db	1390	GATACTGCTACATGGAAATCCAGGTTTGAGAACTGTAAACAGAGCATTACTTGACATGATT	1449
Qy	448	GluGluLysThrValArgAspLysGluLeuGluGlyLeuGlnValLysIleGlnArgLeu	467
Db	1450	GAAGAGAAAGCCATGAGGACCAAGGAATACGAGTGTCTTGTCTGTAATCCAGAGCTA	1509
Qy	468	GluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgValGlnAsp	487
Db	1510	GAGAACTTTGCCGAGCTCTGCAGGAAGAGAGGAATGGAATTGTACAGAAAAATAAACAA	1569
Qy	488	Leu-----SerAlaGlyGlyGlnGlySerLeu---ThrAspSerGlyProGlu	502
Db	1570	GCACAGCTCCCTGAACAGGTGAATGGAATGACATCTTAGAAGAGACGACGATGCCAAT	1629
Qy	503	ArgArgProGluGly-ProGlyAlaGlnAlaProSerSerProA ^{Arg} ValThrGluAlaPro	522
Db	1630	ACAAACCTTCTCTCTCCGAGGAGCGACGATGTGCTGTGCTGACAGAACAATG	1689
Qy	522	cOys-----TyrProGlyAlaProSerThrGluAlaSerGlu	534
Db	1690	CTCAGAGAGCTAGCTGAAGCTTTCAGGTATCCC-----ACAA	1728
Qy	534	yGlnThrGlyProGlnGluProThrSerAla	544
Db	1729	GCAGAGAGAGACCTCCCAAGCGACGCGAGCA	1759
RESULT 32			
AF422244			
LOCUS	2435 bp	mRNA	linear
DEFINITION	Mus musculus muscle-derived protein MDP77 variant 1 (MDp77) mRNA,		ROD 24-JAN-2002
ACCESSION	AF422244		
VERSION	AF422244.1	GI:17026373	
KEYWORDS			
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	1. (bases 1 to 2435)		
TITLE	Benson, K.F. and Chada, K.		
	Molecular characterization of the mouse In(10)17Rk inversion and		
	identification of a novel muscle-specific gene at the proximal		
	breakpoint		
JOURNAL	Genetics 160 (1), 279-287 (2002)		
MEDLINE	21663722		
PUBMED	11805063		
REFERENCE	2 (bases 1 to 2435)		
AUTHORS	Benson, K.F.		
TITLE	Direct Submission		
JOURNAL	Submitted (21-SEP-2001) Medical Genetics, University of Washington,		
FEATURES	1705 Pacific Street NE, Seattle, WA 98195, USA		
source	Location/Qualifiers		
	1..2435		

FEATURES
source

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/chromosome="10"
/map="between Chr1-rs2 and Wmap7"
/sex="female"
/tissue_type="heart"
/dev_stage="adult"
1. 2435
/gene="Mdp77"
157..221
CDS
/gene="Mdp77"
/note="muscle-specific; similar to Gallus gallus
muscle-derived protein encoded by GenBank Accession Number
D89999"
/codon_start=1
/product="muscle-derived protein MDP77 variant 1"
/protein_id="AAL33909.1"
/db_xref="GI:17026374"
/translation="MEINHPDQLSVEHPTPGDSSSNQNGPKQDGERSTGQAPAE
QEGSLHPEKGAHVAEELSRLQEDIIITYGSAAPSEGDASETKQPPNTEAPDNEDV
DYETTBEDIREPTAPDEPAKPEVSNKEKLEKILKGLGKEANLIMONLNKLOAP
BEKLDLFKPYTELLDHRTEOKKLLKQAOCTQREKQDOLOQSEHNRAVLARSKLES
LCRELQRHNTLKESETLQRAEEBEKEITSHFTTLTDLQIQEQSSRNKMLQOE
NTLAEKLSIIIOVELREHLDKIFKHRELQCKLVDAKLEAEQELMCEAEERHREK
BYLNQAEWLQAKVLEKQETVLQAOITLYSGFPEFQSTLTLSNEVFVTFKQEMDK
TTKMKLEKDTATWKFENCNKALLDIMEEKLRAKEVECFVMKFORLENLCRALQ
BERELYKIRIEKSEKEDQVORTSEEPSESVSENEVDABEANSFQKAVENLATA
FTILHPEFTPDQETPQVLPQSGSDVTHQHPETARLNHPSLPADSGSPRPVGA
QAVAEVCEATPATASCTPAEALQSGDPAENTPGKPKHKEPANTSGOAPLSPAGG
SLSVVEAKYDISPSESGDSAVVPGCESRQPPPEVDIFVPGSTGLPREPDACLNG
VD"

```

ORIGIN

Alignment Scores:

```

Pred. No.: 9,29e-50 Length: 2435
Score: 1229,50 Matches: 265
Percent Similarity: 63.89% Conservative: 80
Best Local Similarity: 49.07% Mismatches: 132
Query Match: 44.15% Indels: 63
DB: 10 Gaps: 8

```

US-10-023-523-44 (1-546) x AF422244 (1-2435)

```

Qy 34 ProSerGlnAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArg 53
Db ||| : : : : : |||
Qy 202 CCCCAGGGAGCAGCTCACTCAATCAAAACGGCCCGGAAAGCAAGATGGCGAGCGG 261
Db ||| : : : : : |||
Qy 54 LysProGluGlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeu----- 68
Db ||| : : : : : |||
Qy 262 TGCTCAACCTCAGGCCAAGCG--CCAGAGCAAGAGGGAGCCTGCATCCGAGAGGGA 318
Db ||| : : : : : |||
Qy 69 ---ArgAspValSerGluGluLeuSerArgGlnLeuGluAspIleLeuSerThyCys 87
Db ||| : : : : : |||
Qy 319 GCCCATGATGTCGGGAGAGAGTTCAGCAGGCACTGGAAGACATCATAAAGCAGCTAT--- 375
Db ||| : : : : : |||
Qy 88 ValAspAsnAsnGlnGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluPro 107
Db ||| : : : : : |||
Qy 376 -----GGGCTGCTGCCAGTCCC 393
Db ||| : : : : : |||
Qy 108 GluAspAlaGluLysSerArgThrTyValAlaArgAsnGlyGluProGluProThrPro 127
Db ||| : : : : : |||
Qy 394 GAGGAAGGAGACGCTCTGAAACT----- 417
Db ||| : : : : : |||
Qy 128 ValValAsnGluGluLysGluProSerIlyAspProAsnThrGluGluIle----- 145
Db ||| : : : : : |||
Qy 418 -----AAGGAGACGCCCCCAACACAGAGGCACCAATGAGGATGTTGACTAT 468
Db ||| : : : : : |||
Qy 146 ArgGlnSerAspGluValGlyAspArgAspHisArgArgProGln----- 160
Db ||| : : : : : |||
Qy 469 GAAGAAACCACTGAAGAGATAGACAGAGAACCCCACTGCTCTGAGAGCGACCGCAGCC 528
Db ||| : : : : : |||

```

```

Qy 161 -----GluLysLysLysAlaLysGlyLeuGly 169
Db ||| : : : : : |||
Qy 529 AAAGAGCCTGTCCAGCAATAAAGAGCAAAAGCTGGAAGAAAGAAATCTTAAAGGATTAGGT 588
Db ||| : : : : : |||
Qy 170 LysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeu 189
Db ||| : : : : : |||
Qy 589 AAAGAAGCTAACTTGTCTCATGCAAAACCTGAAACAGCTCAAGACCTGAAGAAAGCTT 648
Db ||| : : : : : |||
Qy 190 AlaAlaLeuCysLysLysTyValaGluLeuLeuGluGluHisArgAsnSerGlnLysGln 209
Db ||| : : : : : |||
Qy 649 GATTTTATTATCAAGAGATATCTGATTTGCTGGATGAACATCGCACCGAGCAAAAGAAA 708
Db ||| : : : : : |||
Qy 210 MetLysLeuGlnLysLysGlnSerGlnLeuValGlnGlnLysAspHisLeuArgGly 229
Db ||| : : : : : |||
Qy 709 CTGAAGCTCTCTACTGAACAGAGCGCGAGCCAGCCAGAGAGAGAGACCATGTTGCAGGT 768
Db ||| : : : : : |||
Qy 230 GluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGln 249
Db ||| : : : : : |||
Qy 769 GAGCAACAAGACCGCTCTTCTGCGAAGCAAACTCGAGAGTCTGTGTCGGAGCTACAG 828
Db ||| : : : : : |||
Qy 250 ArgHisAsnArgSerLeuLysGluGlyValGlnArgAlaArgGluGluGluLys 269
Db ||| : : : : : |||
Qy 829 AGGCACAACAAGACCGCTGAAGGAAGAAACCCCTCAACGGCGCGCGAGGAGAGAGAAG 888
Db ||| : : : : : |||
Qy 270 ArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGlu 289
Db ||| : : : : : |||
Qy 889 AGAAAGAGATCAACAGTCACATTTCAACTACCTCAGAGATATCCAGACTCAGATCGAA 948
Db ||| : : : : : |||
Qy 290 GlnHisAsnGluArgAsnSerLysLeuArgGlnGlnAsnMetGluLeuAlaGluArgLeu 309
Db ||| : : : : : |||
Qy 949 CACACAGTTCAGCGCAATATGAGCTCTGCCAGGAGAACACAGAGCTTCGAGAAAACCTG 1008
Db ||| : : : : : |||
Qy 310 LysLysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAspLysValPheLysHis 329
Db ||| : : : : : |||
Qy 1009 AAAAGCATCATTCAGCAGTACAGAGCTCAGAGAGGAGCATCTGACAAAATATTCAAACAT 1068
Db ||| : : : : : |||
Qy 330 LysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGlnMetLeuLys 349
Db ||| : : : : : |||
Qy 1069 CGAGAACTCGAGCAAGCTGTGATGCGAAACTTGAAGAGCTCAAGAGCTATGATGCGAG 1128
Db ||| : : : : : |||
Qy 350 GluAlaGluArgHisGlnArgGluLysAspPheLeuLysGluAlaValGluSer 369
Db ||| : : : : : |||
Qy 1129 GAGGCAGAGAGCGGCACAGCGGAGAGGAATATTTCTGAAACCGAGCGGAGAGTGG 1188
Db ||| : : : : : |||
Qy 370 GlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeu 389
Db ||| : : : : : |||
Qy 1189 AAACCTTCAGCCAAAGTGTGAGGAGCAGGAGAGCGTCTGCAGGCTCAGCTCACTCTC 1248
Db ||| : : : : : |||
Qy 390 TyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluValPheThr 409
Db ||| : : : : : |||
Qy 1249 TACTCAGGAAGGTTTCAGGAGTTCAGAGTACACTGACAAAAAGCAATGAAGTGTTC 1308
Db ||| : : : : : |||
Qy 410 ThrPheLysGlnGlnMetGluLysMetThrLysLysIleLysLysLeuLysGluThr 429
Db ||| : : : : : |||
Qy 1309 ACTTTCAGCAGGAAATGGAATAGCACTAGAGAAATGAAGAAGCTGGAGAGGACAC 1368
Db ||| : : : : : |||
Qy 430 ThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeuGluMetAlaGluGlu 449
Db ||| : : : : : |||
Qy 1369 GCCACGTGAAGCCGATTCAGATTTGATTAACAAGCCCTGTTGGACATGATCGAGAG 1428
Db ||| : : : : : |||
Qy 450 LysThrValArgAspLysGluLeuGluGlyLeuGlnValLysIleGlnArgLeuGluLys 469
Db ||| : : : : : |||
Qy 1429 AAAGCCCTCGAGCTAAAGATACGAGTGTTCGTGATGAAATCCAAAGGCTGGAGAAC 1488
Db ||| : : : : : |||
Qy 470 LeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgValGlnAspLeuSer 489
Db ||| : : : : : |||
Qy 1489 CTGTGTGTCGTTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1548
Db ||| : : : : : |||
Qy 490 AlaGlyGlyGlnGlySerLeuThrAspSerGlyProGluArgArgProGluGlyProGly 509
Db ||| : : : : : |||
Qy 1549 ATGTCGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1605
Db ||| : : : : : |||
Qy 510 AlaGlnAlaProSerSerProArgValThrGluAlaProCysTyProGlyAlaProSer 529
Db ||| : : : : : |||

```

```

Db      1606 GTCTCTGAAATGAAGAGGTGGACGAGGAGCAACAGCTTTTCAAAAAGCTGTGGAA 1665
Qy      530 ThrGluAlaSerGlyGlnThr-----GlyProGlnGluProThr 542
Db      1666 AACCTGGCCACAGCTTTTACCATCTTCATCATCCAGAGATTCCACCCGACCAACCAACG 1725

RESULT 33
AF422245 LOCUS linear ROD 24-JAN-2002
DEFINITION Mus musculus muscle-derived protein MDP77 variant 2 (Mdp77) mRNA,
complete cds.
ACCESSION AF422245
VERSION AF422245.1 GI:17026375
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 4394)
AUTHORS Benson,K.F. and Chada,K.
TITLE Molecular characterization of the mouse In(10)17Rk inversion and
identification of a novel muscle-specific gene at the proximal
breakpoint
Genetics 160 (1), 279-287 (2002)
JOURNAL Genetics 160 (1), 279-287 (2002)
MEDLINE 21663722
PUBMED 11805063
REFERENCE 2 (bases 1 to 4394)
AUTHORS Benson,K.F.
Direct Submission
Submitted (21-SEP-2001) Medical Genetics, University of Washington,
1705 Pacific Street NE, Seattle, WA 98195, USA
LOCATION/NEIGHBORS
FEATURES
Source
1..4394
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/chromosome="10"
/map="between Cni-rs2 and Mtap7"
/sex="female"
/tissue_type="heart"
/dev_stage="adult"
1..4394
/gene="Mdp77"
157..2211
/gene="Mdp77"
/note="muscle-specific; similar to Gallus gallus
muscle-derived protein encoded by GenBank Accession Number
D89939"
/codon_start=1
/product="muscle-derived protein MDP77 variant 2"
/protein_id="P33910.1"
/db_xref="GI:17026376"
/translation="MEINHPQLSVHEPTPPGDSLSNONGPGKQDGERSTSGQAPE
QEGSHPKGADVAFELSRLEDIIISTYGSAASPEGDASETKEQPNTTEAPNEDV
DYEETIEDIPRPAPEPAKAPKPVSNKEQKSLKILKGLKEANLLMNLKLOAP
EKLDFLPKTYLDEDETRQKLLKLLKQQAQTOREKQDQSEHNRAVLARKLS
LCRELQRNKLKEETLQREBEERKEITSHFTOTLTDIQTQEQSEKMKLQGE
NTELAELKLSIIDQVELREHLDKIFHRELQKLVDAKLEAEQELMOEAERHREK
EYLLNQAWEKLVKQSTVLAQTLTYSGRFEFQSTLTNSVFATFKQEMDK
TTPKMKLEKDTATKAFENCNKALDIMEKALRAKEVECFVYKIKQRLNCRALQ
PERKLYKIRAKMSKEDQVQTSBEPSPVSENEVEADAEANGFOKAVENLATA
FTILHPETDQPTERQLAVNGQSGDVTHQHPETARLNHPSLPADSGSPRPVGA
QVAVGCEATPAPTAPEABVQLQSGDLPAENTPGPKPKPANTSGQAPLSPAG
SUSVVEAKYDISPSESGDSAVVPGCESREQPPPEVTDIIPVGFSTGLFPEDACLNG
VD"

```

ORIGIN

Alignment Scores:
Pred. No.: 1,72e-49 Length: 4394
Score: 1229.50 Matches: 265
Percent Similarity: 63.89% Conservative: 80

```

Best Local Similarity: 49.07% Mismatches: 132
Query Match: 44.15% Indels: 63
DB: 10 Gaps: 8

US-10-023-523-44 (1-546) x AF422245 (1-4394)

Qy      34 ProSerGlnAlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArg 53
Db      202 CCCCAGGGGACAGCTCATCTCAATCAAAACGCGCGGAAAGCAAGATGCCGAGCGG 261
Qy      54 LysProGluGlyValaGlnAlaArgThrAlaGlnSerGlyAlaLeu----- 68
Db      262 TGCTCAACCTCAGGCCAAGCG---CCAGAGCAAGAGGAGGAGCTCATCCCGAGAGGGA 318
Qy      69 ---ArgAspValSerGluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCys 87
Db      319 GCCCATGATGTCGCGGAGAGATTGACAGGCAACTGGAAGACATCATATAAGCACGTAT--- 375
Qy      88 ValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluPro 107
Db      376 -----GGGTCTGCTCCCAAGTCCC 393
Qy      108 GluAspAlaGluLysSerArgThrTyrValAlaArgAsnGlyGluProGluProThrPro 127
Db      394 GAGGAAGGAGAGCGCTCTGAAACT----- 417
Qy      128 ValValAsnGlyGluLysGluProSerLysGlyAspProAsnThrGluGluLeu----- 145
Db      418 -----AAGGAGCAGCGCCCAACACAGAGGAGGAGGAGCAATGAGGATGTTGACTAT 468
Qy      146 ArgGlnSerAspGluValGlyAspArgAspHisArgArgProGln----- 160
Db      469 GAAGAAACCACTGAGAGATAGACAGAGAACCCACTGCTCTCGAGAGCCAGCCGCCAGCC 528
Qy      161 -----GluLysLysLysAlaLysGlyLeuGly 169
Db      529 AAAGAGCCTGTGCAGCAATAAAGAGCAAAAGCTGGAAGAAAGAAATCTTAAAGGATTAGT 588
Qy      170 LysGluLeuThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeu 189
Db      589 AAAGAAGCTTAACCTTGCTCATGCAAAACCTGAAACAGCTGCAAGACCTCGAAGAAAGCTT 648
Qy      190 AlaAlaLeuCysLysLysTyrAlaGluLeuLeuGluHisArgAsnSerGlnLysGln 209
Db      649 GATTTTTTATTCAGAGAGTATCTGTAATGCTGGATGAACATCGCACCGCAAAAGAAA 708
Qy      210 MetLysLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGly 229
Db      709 CTGAAGCTCTCTACTGAACAGCAGCGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 768
Qy      230 GluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGln 249
Db      769 GAGCACACAGAGCGCTCTTCTGCTCGAAGCAAACTCGAGAGTCTGTGTCGGAGCTACAG 828
Qy      250 ArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGluGluGluLys 269
Db      829 AGGCACAACAAGAGCGCTGAAGGAAGAAACCTCCCAACGCGCGCGGAGGAAGAGAGAG 888
Qy      270 ArgLysGluValThrSerHisGlnValThrLeuAsnAspIleGlnLeuGlnMetGlu 289
Db      889 AGGAAAGAGATCAAGTCACTTCAAACTACCTGACAGATATCCAGACTCAGATCGAA 948
Qy      290 GlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeu 309
Db      949 CAACAGAGTGAGCGCAATATGAAGCTCTGCCAGGAGAACACAGAGCTTCGAGAAAACCTG 1008
Qy      310 LysLysLeuIleGlnThrGluLeuArgGluGluHisIleAspLysValPheLysHis 329
Db      1009 AAAAGCATATTGACCAGTACGAGCTCAGAGGAGGAGCATCTGGACAAAATATATCAACAT 1068
Qy      330 LysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLys 349
Db      1069 CGAGAACTGCAGCAGAGAGCTGGTGGATGCGAAACTTGAAGAGAGCTCAAGAACTGATGCAG 1128

```



```

RESULT 35
AR072731 LOCUS AR072731 1191 bp DNA linear PAT 28-AUG-2000
DEFINITION Sequence 5 from patent US 5948643.
ACCESSION AR072731
VERSION AR072731.1 GI:99999495
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 1191)
Rubinfeld,B., Polakis,P.G., Lingenfelter,C. and Vuong,T.T.
AUTHORS Modulators of BRCA1 activity
TITLE Patent: US 5948643-A 5 07-SEP-1999;
JOURNAL Location/Qualifiers
FEATURES
1..1191
source
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1,58e-41 Length: 1191
Score: 1048.50 Matches: 230
Percent Similarity: 69.01% Conservative: 55
Best Local Similarity: 55.69% Mismatches: 104
Query Match: 37.65% Indels: 24
DB: Gaps: 6

US-10-023-523-44 (1-546) x AR072731 (1-1191)
QY 19 SerProGlyGlnProGluAlaGlyProGlu-GlyAlaGlnGluArgProSerGlnAlaAl 38
DB 12 TCCCCCGGGCTGCAGGAATTCGCGACGAGCGCGCGCGAGAG-----54
QY 38 aProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLys-----ProG1 56
DB 55 ----GCACACTGAGCGCGAGCGGGCGGCGCGAGCGCGAGCGCGCGAGAGTTTGAAT 110
QY 56 uGlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuA:9AspValSerGluGluLeu 76
DB 111 TGGCACAATGAAGAAGCTGGAATTTGTGGGCTAGGGGTGAAAGCAGATATGTTGTGTA 170
QY 76 rArgGlnLeuGluAspIleLeuSerThrTyrcysValAspAsnAsnGlnGlyProG1 96
DB 171 CTCTCAATCAATATGATATCTTCAACATCAA-----GGCTCAATGTGTGCGCACAG 224
QY 96 yGluAspGlyAlaGlnGlyProAlaGluProGluAspAlaGluLysSerArgThrTy 116
DB 225 TAACAAGCATTTCATTGGAAGAGGATGAAGGCGAGTCTTTTATAACAGAGAAAGAAATTT 284
QY 116 rValAlaArgAsnGlyGluProGluProThrProValValAsnGlyGluLysGluPro 136
DB 285 GGTGAGC-----CCAGCATCTGCACCGCAAGAATCAAGAGAGGAAT 326
QY 136 rLysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHi 156
DB 327 CCCTGGGGGA-----GAAGCTCGAACAGATCCCTCTGTGTCAGCAAGATTC 374
QY 156 sArgArgProGlnGlnLysLysAlaLysGlyLeuGlyGluIleThrLeuLeuMe 176
DB 375 AGAGTGCACAGGCAACAAAGAAAAAAGT-----TTAGGAAAAAGAGTTTATTACTGAT 428
QY 176 tGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysTy 196
DB 429 GCAAGCCCTAAACACCTTTTCAACCCCGAGAGAGAGTCTGGCAGCTCTCTGTAAAGAAATA 488
QY 196 rAlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuGlnLysLy 216
DB 489 TGCTGATCTTCTGAGGAGAGCAGGAGTGTTCAGAAAGCAAAATGAAGATCTCTGCAGAGAA 548
QY 216 sGlnSerGlnLeuValGlnGlnLysAspHisLeuArgGlyGluHisSerLysAlaVal 236
DB 549 GCAAGCCCAAGATTTGTAAGAGAGAAAGTTTCACTTTCAGAGTGAACATAGCAGGCTATCTT 608

```

```

78 GlnLeuGluAspIleLeuSerThrTyrcysVal---AspAsnAsnGlnGlyGlyProGly 96
DB 318 AACAGAAATTTGGTGAAGCTCAGTATTTCTACACAGGAGAGAAAGAGAAATTTCTCGA 377
QY 97 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTy 116
DB 378 CGAAGAGCTGGAACAGCTCTCTCTGATGCCAGCAGATTCAGAGTGCAGCAGG----- 431
QY 117 ValAlaArgAsnGlyGluProGluProThrProValValAsnGlyGlyLysGluProSer 136
DB 431 -----431
QY 137 LysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHis 156
DB 431 -----431
QY 157 ArgArgProGlnGlnLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMet 176
DB 432 -----AACAAAGAGAGAGCTTAGGAAAGAGAGTTTATTACTGATG 473
QY 177 GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysTy 196
DB 474 CRAAGCGCTAAACACCTTTCAACCCCGAGAGAGAGCTGCGAGCTCTCTGTAAAGAAATAT 533
QY 197 AlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLys 216
DB 534 GCTGATCTCTGGAAGAGAGAGAGATGTTTCAGAAACAAATGAAGATTTCTGCAAGAGAG 593
QY 217 GlnSerGlnLeuValGlnGlnLysAspHisLeuArgGlyGluHisSerLysAlaValLeu 236
DB 594 CRAAGCCAGATTTGTAAGAGAGAAAGTTTCACTTCAGAGTGAACACAGCAAGGCCATCTTG 653
QY 237 AlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLys 256
DB 654 GCAAGAGAGCAACTGGAATCTCTTTTCAGGGAAGTTTCAAGCTGATTAATGAACCTTAAG 713
QY 257 GluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHis 276
DB 714 GAGGAGAAATATGAGCAGGAGCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 773
QY 277 PheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSer 296
DB 774 TTCACGATAACTTAAATGAATCCAGCTCAGTGTGGAACACATGACATCCACATGCC 833
QY 297 LysLeuArgGlnGlnLysMetGluLeuAlaGluArgLeuLysLysLeuIleGlnTy 316
DB 834 AAATCGGACAGAGAGACATTGAACTGGGAGAGAGAGTTGAAGAGCTTATTGAGCAGTAT 893
QY 317 GluLeuArgGluGluHisLysAspLysValPheLysHisLysAspLeuGlnGlnLeu 336
DB 894 GCACCTAAGGAGAGAGATATTGATAAGATTTCAACACACAGAGAAATTCACACACAGCTT 953
QY 337 ValAspAlaLysLeuGlnGlnAlaGlnGlnMetLeuLysGluAlaGluArgHisGln 356
DB 954 GTGATGCGCAACTTCAGCAACACACAGCTGATTAAGAGAGCTGATGAAAAACATCAG 1013
QY 357 ArgGluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMet 376
DB 1014 AGAGAGAGAGAGTTTATTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1073
QY 377 LysGlnGlnGlnThrHisLysLysGlnGlnLeuAlaLeuTyThrGluLysPheGluGlu 396
DB 1074 AACACAGAGAGAGATCAACTAAACAGAGAGCTTCTCTTATATGATTAATTTAGAGAA 1133
QY 397 PheGlnAsnThrLeuSerLysSerSerGluValPheThrPheLysGlnGlnMetGlu 416
DB 1134 TTCACGACTACTATGCAAAAGCAATGAATTTTACAACTTTCAGGAGAGAGAGAGAGAG 1193
QY 417 LysMetThrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 427
DB 1194 AAGATGACAAAGAAATTTAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1226

```



```

/db_xref="SPTREMBL:Q8N3S2"
/translation="OKKLUKQKQVQIQKQKQDQLOGSHSRAILLARSKLSELCRELOR
HNKTLKALQARREBEERKEITSHFQSTLTDIQGIEQSSERNMEJCQNTIELAEK
LAKSIIDQVELREHLDKIFHRELOQKLVDAKLEQAQEMKEAEERHREKEYLLNQ
AWKLOAKMLKEQTVLOQLTLYSGREFEFQSTTKGNEVFATFKQMDKTKMKK
LEKDTATKAPENCNKNALLDMEBKARAKYECFVMKIGRLENLCALQOERNEHL
KIRDAEISEKDDQSHNSDEPEPNSVDQIEIDAEVNSVCTAVKNLATAFMI1HPH
ESTPHQKSTPEIGSSQESADALKEPEQPLIPSRDSEPLPLTPQASAEAGSDA
BPPSKASNPAGLGAETOCCEPLFVGADQDQPSWKEAEASGQAPORPTEASIQKWEAD
VPAPACAAEEHVAAVMVPAPEPSRPPRAAAEELPVGASAGPOPNRVADTNLEGVD"
3957..3962
/gene="DKFp451A175"
polyA_signal
3979
polyA_site
/gene="DKFp451A175"

ORIGIN
Alignment Scores:
Pred. No.: 4,75e-39 Length: 4154
Score: 1008.00 Matches: 203
Percent Similarity: 67.09% Conservative: 64
Best Local Similarity: 51.01% Mismatches: 71
Query Match: 36.19% Indels: 60
DB: 9 Gaps: 2

US-10-023-523-44 (1-546) x HSM803629 (1-4154)

QY 207 GlnLysGlnMetLysLeuLysGlnSerGlnLeuValGlnGlnLysAspHis 226
Db 1 CAAAGAAGTTAAAGTCTCTCCAAAAGAACAGGTACAAATTCAAAAGAAAGACAG 60
QY 227 LeuArgGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLysArg 246
Db 61 TTACAAAGTGAACACAGACAGAGCTATCTCGTCCGAAGCAATTCAGAGTCTGTCCGG 120
QY 247 GluLeuGlnArgHisAsnArgSerLysGluGluValGlnArgAlaArgGlu 266
Db 121 GAGCTCGAGACACAAACAGACTCTGAAGGAGAGGCGCTTCAGCGGACAGTGAAGAA 180
QY 267 GluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspLeu 286
Db 181 GAAGAGAAAGAGAGAAATCAACACCCATTCACAGAGTACCTCCAGCAGATCCAGGGC 240
QY 287 GlnMetGluGlnHisAsnGlnArgAsnSerLysLeuArgGlnGlnAsnMetGlu 306
Db 241 CAGATCGAGCAGCAGAGTGCAGCGAAATATGGAGCTCTGTCCAGAGAACACAGAGCTTGA 300
QY 307 GluArgLeuLysLysLeuLeuGluGlnTyrGluLeuArgGluGluHisLysVal 326
Db 301 GAAAGCTGAAAAGCAATCATCGATCATGATGAGTTCAGAGAGGAGCATCTGGACAAAATA 360
QY 327 PheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnGlu 346
Db 361 TTTAAACACAGAGAACTGCAGCAGAAAGCTGGTGGATGCAAGCTTCAGAGGCCCAAGAA 420
QY 347 MetLeuLysGluAlaGluArgHisGlnArgGluLysAspPheLeuLysGluAla 366
Db 421 ATGATGAAGGAGCGGAGGAGGACACAAACGAGAAAGAAATATTTGCTGACACAGGCA 480
QY 367 ValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlnThrHisLysLysGln 386
Db 481 GCAGAGTGGAAACTTCAGGCGGAAATGCTGAAGGAGCAAGAGACAGATCTCTGAGGCTCAG 540
QY 387 LeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGlu 406
Db 541 CTCACCTCTACTCAGGAGGTTGAAGAAATTCAGAGACACACTAACTAAACACAGGAG 600
QY 407 ValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLysLeuGlu 426
Db 601 GTGTTTCCACGTTCAACACAGGAAATGGACAAACAACTAAGAAATGAAGAGCTGGAA 660
QY 427 LysGluThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeuGluMet 446
Db 661 AAGGACACAGCCACATGGAAAGCCCGATTTGAGAACTGTAAACAAAGCTCTGTGTGGACATG 720

```

```

336 rGluLeuArgGluGluHisLysAspLysValPheLysHisLysAspLeuGlnGlnLe 336
849 GCACCTGAGGAGAGACACATTTGATAGTGTTCACACATAAGCACTGCAACAGCAGCT 908
QY 336 uValAspAlaLysLeuGlnGlnAlaGlnGlnMetLeuLysGluAlaGluGluArgHisGlu 356
Db 909 CGTGATGCCAAACTGCAGCAACACACAACTGATGATAAGAGCTGATGATAACATCA 968
QY 356 nArgGluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMe 376
Db 969 GAGAGAGAGAGAGTTTATTAAAGAGACGACAGATCGAGGACAAATACGACAAAT 1028
QY 376 tLysGlnGlnGlnThrHisLysLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGlu 396
Db 1029 GAAACAGCAAGAGTACAACTAAACACAGCAGCTTCTCTTATATGGATAAGTTGAAGA 1088
QY 396 uPheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGlu 416
Db 1089 ATTCCAGACTACCATGGCAAAAGCAATGAACTGTTTACAACTTCAGACAGGAAATGGA 1148
QY 416 uLysMetThrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 428
Db 1149 AAGATGACAAAGAAATTAATAAAAAAAAAAAAAAAAAAAAAA 1185

HSM803629 4154 bp mRNA linear PRI 13-MAY-2003
LOCUS Homo sapiens mRNA; cDNA DKFp451A175 (from clone DKFp451A175).
DEFINITION AL832322
ACCESSION AL832322.1 GI:21732875
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 4154)
Koehler, K., Beyer, A., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
Fobo, G., Han, M. and Wiemann, S.
Direct Submission
Submitted (13-MAY-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuberberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkz-heidelberg.de;
sequenced by EMFZ (Biomedical Research Center at the
Heinrich-Heine-University, Dueseldorff/Germany) within the cDNA
sequencing consortium of the German Genome Project. This clone
(DKFp451A175) is available at the R2PD in Berlin. Please contact
the R2PD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@r2pd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
Location/Qualifiers
1..4154
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="R2PD:DKFp451A175"
/db_xref="taxon:9606"
/chromosome="6"
/clone="DKFp451A175"
/tissue_type="human skeletal muscle"
/clone_lib="451 (synonym: hiccl). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
1..4154
/gene="DKFp451A175"
<1..1518
/gene="DKFp451A175"
/notes="similar to chicken 77K muscle-derived protein"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAD38617.1"
/db_xref="GI:21732876"

```

FEATURES
source

gene
CDS


```

QY 447 AlaGluGluYsThrValArgAspLysGluLeuGluGlyLeuGlnValLysIleGlnArg 466
Db 721 ATTGAAGAGAAACATCGAGCTAAGAAATATGCTGTTGTGATGAATTCGGGAGG 780
QY 467 LeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgValGln 486
Db 781 CTAGAGAACTCTGCGGTCTTTACAGAGAGAGAGAAACGAACCTCCACAAAATATCAGA 840
QY 487 AspLeuSerAlaGlyGlnGlySerLeuThrAspSerGlyProGluArgProGlu 506
Db 841 GAGCGAGAATATCTGAAAGGATGACCAAGTTCAGCACAACTCCGATGAAGAGCCAGAG 900
QY 507 Gly----- 507
Db 901 TCAACGCTCTCTGTGATCAAGAGATTGACGAGAGAGGTTAATAGTGTCAAACGCC 960
QY 507 ----- 507
Db 961 GTGAAAATCTGGCCACAGCCTTCATGATAATTCATCCAGAGTCACCCGCCAGCAG 1020
QY 508 -----ProGlyAlaGln 511
Db 1021 TCCAAAGAAACCAACCCGAAATAGCGAGTCTTCAGGAGAGTGCTGACGCCCTCTCAAA 1080
QY 512 AlaProSerSerPro-----ArgValThrGluAlaProCysTyrProGly 526
Db 1081 GAGCCAGAGCAACCCCTCTGATCCCTTCACGGGATTCAGAGAGTCCCTGCTCCCTCA 1140
QY 527 AlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAla 544
Db 1141 ACTCTCAGGCTGAAGCCGAGGAGCGAGTGTGCTGAACCTCCCTCCAGGCC 1194

RESULT 38
AX053282
LOCUS AX053282 823 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 48 from Patent WO0073801.
ACCESSION AX053282
VERSION AX053282.1 GI:12227601
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Obata, Y.
TITLE Breast, gastric and prostate cancer associated antigens and uses
thereof
JOURNAL Patent: WO 0073801-A 48 07-DEC-2000;
LUDWIG INSTITUTE FOR CANCER RESEARCH (US)
FEATURES
    source
        1..823
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 6,98e-36 Length: 823
Score: 925.50 Matches: 197
Percent Similarity: 86.55% Conservative: 41
Best Local Similarity: 71.64% Mismatches: 37
Query Match: 33.23% Indels: 5
DB: 6 Gaps: 0

US-10-023-523-44 (1-546) x AX053282 (1-823)

QY 179 LeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysThrAlaGlu 198
Db 2 CTAAACACCCCTTCAACCCAGAGAGAGTGGCAGCTCTCTGTAAGAAATGCTGAT 61
QY 199 LeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLysLeuGlnLysSer 218

```

```

Db 62 CTTCTGGAGAGACGAGAGTGTTCAAGAACAAATGAAGATCTCTGCAGAAAGAACAGCC 121
QY 219 GlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArg 238
Db 122 CAGATTGTGAAGAGAAAGTTCACTTCAGAGTGAACATACAGAGGCTATCTTGGCAAGA 181
QY 239 SerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGlu 258
Db 182 AGCAAGCTTAGAATCTCTTTTCAGAGAACTTCAGCGTCAACAATAAGAGCTTAAAGAGGAA 241
QY 259 GlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGln 278
Db 242 AATATGAGCAGGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 301
QY 279 ValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeu 298
Db 302 ATTACCTTAAATGAATTCAGCCAGCTGGAGCAGCATGACATCCCAACAGCCCAACTC 361
QY 299 ArgGlnGlnAsnMetGluLeuAlaGluArgLysLysLeuLysLeuLysGlnThrGluLeu 318
Db 362 CGACAGAGAAACATTGAGCTGGGAGAGAGCTTAAAGAGCTCATCGAACAGTAGCACTG 421
QY 319 ArgGluGlnHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeuValAsp 338
Db 422 AGGGAAGAGACATTTGATTAAGTGTTCAAACATAAGGAACCTGCAACAGCAGCTCGTGGAT 481
QY 339 AlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGlnArgGlu 358
Db 482 GCCAACTGTCAGCAACACGACCACTGATATAAGAGAGCTGATGAAAACATCAGAGAG 541
QY 359 LysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGln 378
Db 542 AGAGAGTTTATTAAAGAGCGCAGAGATCGAGGCACAAATACGACAAATGAACAG 601
QY 379 GlnGluThrHisLeuLysGlnGlnLeuAlaLeuThrThrGluLysPheGluGluPheGln 398
Db 602 CAAGAGATCACATA-AAACAGCAGCTTCTCTTATATGATAGTAAAGTTGAAGATTTCAG 660
QY 399 AsnThrLeuSerLysSerGluValPheThrThrPhelLysGlnGluMetGluLysMet 418
Db 661 ACTACCATGCAAAAGCAATGACTGTTTCA-ACCTTCAGACAGGAATGGGAAGAAT 719
QY 419 ThrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 438
Db 720 GCCAAGAAAATTAA-AAACTGGGAAAAGAAACCACTTAATTTGGGGTTCCTCAATGGGAAAC 778
QY 439 SerAsnLysAlaLeuLeuGluMetAlaGluLysThrValArg 453
Db 779 AATATTAACN-CTTTTGCA-ATGGNTGAAGAAAAACAGTCCCG 821

RESULT 39
AX834462
LOCUS AX834462 2194 bp DNA linear PAT 15-DEC-2003
DEFINITION Sequence 1586 from Patent EP1347046.
ACCESSION AX834462
VERSION AX834462.1 GI:39920597
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S.,
Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R.,
Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and
Masuko, Y.
TITLE Full-length cDNA sequences
JOURNAL Patent: EP 1347046-A 1586 24-SEP-2003;
Research Association for Biotechnology (JP)
FEATURES
    source
        1..2194
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"

```

/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 2,12e-34 Length: 2194
 Score: 903.50 Matches: 257
 Percent Similarity: 47.28% Conservative: 38
 Best Local Similarity: 41.19% Mismatches: 112
 Query Match: 32.44% Indels: 217
 DB: 6 Gaps: 18

US-10-023-523-44 (1-546) x AX834462 (1-2194)

QY 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerPro 20
 Db 66 ATGAAGAACCAAGACAAAGAACCGGGCTGCCAAATCCAAATCCAAAGACGCCCA 125
 QY 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
 Db 126 GGACACCGAAGCAGGACCGAGGAGCCAGAGCGGCCCGCCAGCGCGGCTCCTGCA 185
 QY 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
 Db 186 GTAGAAGCAGGAGGTCCCGGCGAGCAGCCAGGCTCTCGGAAGCCGAGGCGGCTCAAGCC 245
 QY 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
 Db 246 AGAAGCGGTCTAGTCGGGGCCCTTCGTGATGTCTCTGAGGAGTGTAGCGCGCCCACTGGAA 305
 QY 81 AspIleLeuSerThrTyrcysValAspAsnAsnGlnGlyProGlyGluAspGlyAla 100
 Db 306 GACATCTAGCACATCTGTGTGGAATAACAGCGGGGGCCCGCGGAGGATGGGCA 365
 QY 101 GlnGlyGluProAlaGluProGluAspAla-----Glu-LysSerArgTh 115
 Db 366 CAGGTGAGCGCGGTGAACCGCAAGATCTGCCCTCTTAGTTCATAGGCCAAAGTGATGA 425
 QY 115 rTyValAlaArg-----AsnGlyGluProGluProThrProValAlaAsn----- 130
 Db 426 TCGTGTGTGAGGACCTAGAGGGCGCTCCCTGACCCACCCCTTCTCTTGCCTACTTCA 485
 QY 131 -----GlyGlyLys-----GluProSerLysGlyAspProAsnThrGl 143
 Db 486 TCCTCTGGGACAAAGCTGCTGTGTGTGGTGGAGGAGTGTGTGTGTCTTATCCCTC 545
 QY 143 uGluIleArgGlnSerAspGluValGlyAspArgAspHisArgArgProGlnGlu----- 161
 Db 546 AGCGCTGAGACATAGAGGCTTCCTGGGCCACTACAGTGAGACACGAACCTCAAGAATCTG 605
 QY 162 -----LysLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLe 175
 Db 606 AATACCCCGGTTTCTCTCCCGCCAGGCAAAA-----AAGGACTTAGTACTACC 656
 QY 175 uMetGln-----ThrLeuAsnThrLeuSerThrProGl 186
 Db 657 TGTGGAGAAGAGGTGCAGGACTACCAGGCCCTGTCTTGTGATTTACAGCCCTCCCA 716
 QY 186 uGluLysLeuAlaAlaLeuCysLysLysTyAlaGluLeuGluGluHis----- 203
 Db 717 GACAGACAGGACCCCTCATCATACCCAACTGGGACTTACCTGCTAGGCACCTTCCCTT 776
 QY 204 -----ArgAsn----- 205
 Db 777 CCCCATCCAAAAAATGGAGTTATTTTCCCTTATTTTCAAGAGTCCAGTTGATTTTACCT 836
 QY 206 -----SerGlnLysGlnMetLysLe 212
 Db 837 TTGAAGTAGACCTGAGTCTCTTCCATCCCTTCTCTCTCACCTGACACAGG 896
 QY 212 uLeuGlnLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHis 232
 Db 897 TCTGACGCGCTCCTCTAGT-----AGGACGAGCAGCCATCTCTTGGGGATGCACAT 947

QY 232 rLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArg----- 246
 Db 948 GTCTAGTCTTTTGGCTA-----CATATGCAAGTCTTTGCCCACTGAGCTAGCTGTTAT 1001
 QY 247 -----G1 247
 Db 1002 GTTCTTAGAGGCATTGTTTGGCCATTCTTCCATTACAGAGAATCAGGACACAGA 1061
 QY 247 uLeuGlnArgHisAsnArgSerLeuLysGluGluGlyValGln----- 261
 Db 1062 AGTGAGGCTTCCAGCCCATAGGTATCAATCTGGGGTCAGAGATTGAGTGTGTTAT 1121
 QY 262 -----ArgAlaArgG1 265
 Db 1122 TGCTTGCCTTTTGGGAGCAGATTCCATCCATAAACCATGTCTTACCAAGGTGACTC 1181
 QY 265 uGluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleG1 285
 Db 1182 ACTGGGAGAGAAACGACGTGAGTGGGA-----AAGCTGACCTTCCAGAGACTTGG 1232
 QY 285 nLeuGlnMetGluGlnHis-----AsnGluArgAsnSerLysLeuArg 299
 Db 1233 GGCCCATGTTGTGTGTGTCACATGGGAGTCCATCATATCAGATTGAGATGGGGCTGG 1292
 QY 299 gGlnGluAsnMetGluLeuAla-----GluArgLeuLysLe 312
 Db 1293 CAAAGTGCCTGTCTGTGTGTGGGTACCTGAGAAAGGGAGCGCTGACAGCCG 1352
 QY 312 uIleGluGlnTyArgGluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLe 332
 Db 1353 ACTGCTCCCAACCATCTTTGTTG---CAGCATATGCACAAAGTCTTCAACACAAAGGAC 1409
 QY 332 uGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnLeuLysGluAlaG1 352
 Db 1410 ACACAGCAGCTGTGTGGATGCCAAGCTCCAGAGCCAGGAGATGCTAAAGAGGAGCAGA 1469
 QY 352 uGluArgHisGlnArgGluLysAspPhe----- 361
 Db 1470 AGAGCGGCACACAGCGGAGAGAGATTTTGTGAGGCTCAGGCCCCAGGGTTGGGGTGGGG 1529
 QY 361 ----- 361
 Db 1530 TGGGAGGACAGGCTGGGCTCTGCTCAGTCTAGTCAGCGGGTTATATGGAGAAAGCTG 1589
 QY 361 ----- 361
 Db 1590 GCCAGACAGGACAGATTCTCTTGTAGTACCAGTCTGAGAGCAGGAAGCCTCAGTGGGTCT 1649
 QY 361 ----- 361
 Db 1650 GGTGCTTGTGGCTAAACAAACATAGCCCTGGGGCTTCTGACAGGATCTGGGGTTC 1709
 QY 362 -----LeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetL 377
 Db 1710 TGTCTTGGAAATAGTCTCTGAAAGAGGAGTAGAGTCCAGAGGATGTGTGAGCTGATGA 1769
 QY 377 ysgGlnGlnThrHisLeuLysGlnGlnLeuAlaLeuTyThrGluLysPheGluGluP 397
 Db 1770 AGCAGCAGAGACCCACCTGAAAGCAACAGCTTGCCTTATACACAGAAAGTTGAGGAGT 1829
 QY 397 heGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluL 417
 Db 1830 TCAGAAACACATCTTCCAAAGCAGCGAGGTATTCCACCATTCACAGCAGGAGATGGA 1889
 QY 417 ysmetThr 419
 Db 1890 AGGTAACT 1897

RESULT 40
 AK096939
 LOCUS
 DEFINITION Homo sapiens cDNA FLJ39620 fis, clone SMINT200114.
 ACCESSION AK096939

PRI 15-JUL-2002

```

VERSION      AK096939.1 GI:21756551
KEYWORDS     oligo capping; fls (full insert sequence).
SOURCE       Homo sapiens
ORGANISM     Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE    1
AUTHORS      Kanehori, K., Ishibashi, T., Chiba, Y., Fujimori, K., Hirao, S.,
              Tanai, H., Watanabe, S., Ishida, S., Ono, Y., Hotuta, T., Watanabe, M.,
              Sugiyama, T., Irie, R., Osuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
              Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,
              Kimura, K., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,
              Wagatsuma, M., Takahashi-Fujii, A., Oshima, A., Sugiyama, A.,
              Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuho, Y., Nagai, K.,
              and Isogai, T.
TITLE        NEDO human cDNA sequencing project
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 2194)
AUTHORS      Isogai, T. and Yamamoto, J.
TITLE        Direct Submission
JOURNAL      Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
              Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
              (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
COMMENT      NEDO human cDNA sequencing project supported by Ministry of
              Economy, Trade and Industry of Japan; cDNA full insert sequencing:
              Research Association for Biotechnology (RAB); cDNA library
              construction: Helix Research Institute (HRI) (supported by Japan
              Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
              HRI, and Biotechnology Center, National Institute of Technology and
              Evaluation; clone selection for full insert sequencing: HRI and
              RAB; annotation: HRI and RAB.
FEATURES     Location/Qualifiers
              1..2194
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="SMINT2001114"
                /tissue_type="small intestine"
                /clone_lib="SMINT2"
                /note="cloning vector: pME18SFL3"
ORIGIN
source
Alignment Scores:
Pred. No.:      2,128-34      Length:      2194
Score:          903.50      Matches:    257
Percent Similarity: 47.28%      Conservative: 38
Best Local Similarity: 41.19%      Mismatches: 112
Query Match:     32.44%      Indels:     217
DB:              9          Gaps:         18

US-10-023-523-44 (1-546) x AK096939 (1-2194)

QY 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerPro 20
DB 66 ATGAAGAACCAGACAAAGAAAGACGGGCTGCCAATCAATCCAAAGAGCAGCCCA 125
QY 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
DB 126 GGACACCCGAGACGAGACCCGAGGAGCCAGGAGCCGAGGAGCGGCTCTGCA 185
QY 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
DB 186 GTAGAGACGAGGAGGTCCCGGACGACGACGAGCTCTCTCGAAGCCGAGGGGCTCAAGCC 245
QY 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
DB 246 AGAAGCGGTCACTCTGGGGCCCTTCGTATGCTCTGAGGAGCTGAGCGCCCACTGGAA 305
QY 81 AspIleLeuSerThrTyrcysValaspAsnAsnGlnGlyProGlyGluAspGlyAla 100
DB 306 GACATACCTGAGCACATCTGTGTGACAAATACCAAGGGGGGCCCGCGGAGATGGGGCA 365
QY 101 GlnGlyGluProAlaGluProGluAspAla-----Glu-LysSerArgTh 115

```

```

Db 366 CAGGCTGAGCGGCTGAACCCGAAGATCTCCCTCTTAGTTCATAGCCCAAGTGTATGA 425
QY 115 rTyValAlaArg-----AsnGlyGluProGluProThrProValValAsn----- 130
Db 426 TCGGTGTGCAGGACCTAGAGGGCGCTCCCTGACCCACCCCTTCCTTCCCTCACTTCA 485
QY 131 -----GlyGluLys-----GluProSerLysGlyAspProAsnThrGl 143
Db 486 TCCTCTCGGACAAAGCTGTGTGTGGTTGAGGGAGTTGGTTGGTTCTTCTTATCCCTC 545
QY 143 uGluIleArgGlnSerAspGluValGlyAspArgAspHisArgArgProGlnGlu----- 161
Db 546 AGCCTGAGACATAGAGGCTTCCTGGGCCCACTACAGTGACACACGAACCTTCAAGAATCTG 605
QY 162 -----LysLysLysAlaLysGlyLysGlyLeuGlyLysGluLeuLeuLe 175
Db 606 AATACCCCGGTTTCTCTCCCGCCAAAGCAAAA-----AAGGACTTAGTACTACC 656
QY 175 uMetGln-----ThrLeuAsnThrLeuSerThrProGl 186
Db 657 TGTGGAGAAGAGGTGCAGGACTACACAGGCCCTGCTGCTTTGCAATTTACAGCCCTCCCA 716
QY 186 uGluLysLeuAlaAlaLeuCysLysLysTyraAlaGluLeuGluGluHis----- 203
Db 717 GACAGACACAGGACCCCTCATCATACCAACTGGACTTACCTGCTAGGACCTTCCCTT 776
QY 204 -----ArgAsn----- 205
Db 777 CCCATCCAAAAAATAGGATTTATTTCCCTTATTTCCAGCAAGTCCAGTTGATTTACCT 836
QY 206 -----SetGlnLysGlnMetLysLe 212
Db 837 TTGAAGTAGCACCTGAGTCTCTCACCTTCTCCATCCCTTCTCTCACTGACACAGG 896
QY 212 uLeuGlnLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSe 232
Db 897 TCTCAGCGCTCTCTCTAGT-----AGGCAGACAGCCATTCCTTGGGATGCACAT 947
QY 232 rLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArg----- 246
Db 948 GTCTAGTCTTTGCCCTA-----GATATGGCAAGTCTTTGCCAACTGAGCTAGCTGTAT 1001
QY 247 -----Gln 247
Db 1002 GTTCTTAGAGGCATTTGTTTGGCCATTTCTCCCATTTTACAAGAGATCAGGACACAGA 1061
QY 247 uLeuGlnArgHisAsnArgSerLeuLysGluGluGlyValGln----- 261
Db 1062 AGTGAGGCTTCCAGGCCCATAGGTGATCAATCTCGGGGTTCAGAGATTTAGTGTGTAT 1121
QY 262 -----ArgAlaArgGl 265
Db 1122 TGCTTGCTTCTTGGGAGCAGATTCATCCATAAACCATGTGCTTACCAAGTCTGACTC 1181
QY 265 uGluGluGlyArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGl 285
Db 1182 ACTGGGAGAGAAACGACGTGAGGTTGGA-----AAGCTGACCTTCCAGAGACTGG 1232
QY 285 nLeuGlnMetGluGlnHis-----AsnGluArgAsnSerLysLeuAr 299
Db 1233 GGCCCATGTTGTGTGTGTACATCGGAGTCCATCATATCAGATTGATGGGGGGCTGGG 1292
QY 299 gGlnGluAsnMetGluLeuAla-----GluArgLeuLysLysLe 312
Db 1293 CAAAGTGCCCTGCTGTGTGCTGCGGTACCTCGAGAAAGGGAGCGCGCTGACAAGCCG 1352
QY 312 uIleGluGlnTyrgluLeuArgGluGluHisIleAspLysValPheIshLysAspLe 332
Db 1353 ACTGCTCCCACTCTTTGTG---CAGCATATCGCAAGAGTCTTCAACACAGGACCT 1409
QY 332 uGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMetLysGluAlaGl 352

```

Db 1410 ACAACAGCAGCTGGTGGATGCGCAAGCTCCAGCAGGCCAGGAGATGCTTAAGGAGGCAGA 1469
 QY 352 uGluArgHsGlnArgGluysAspPhe----- 361
 Db 1470 AGAGCGGCACACGCGGAGAGAGATTTTGTGAGGCTCAGGCCCCAGGCTTGGGTGGGG 1529
 QY 361 ----- 361
 Db 1530 TGGGAGGAGACAGGCTGGGCTCTGGCTAGCTCATAGCGGGTTATATGGGAGAGATCTG 1589
 QY 361 ----- 361
 Db 1590 GCCAGACAGGACACAGATTCTTGTAGTACCAGTCTCAGAGCAGGAAGCTCAGTGGGTCT 1649
 QY 361 ----- 361
 Db 1650 GGTGCTTGTGCTTAAACCAACATAGCCCTCGGGGCTTCTGACAGATCTGGGGTTC 1709
 QY 362 -----LeuLeuLysGluAlaValGlnSerGlnArgMetCysGluLeuMetL 377
 Db 1710 TGTCTTGGAAATAGCTCTCTGAAAGAGGAGTGTAGTCCAGAGGATGTGTGAGCTGATGA 1769
 QY 377 ySglnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluP 397
 Db 1770 AGCAGCAGAGAGACCCACCTGAAAGCAACAGCTTGCCCTATACACAGAGAGATTGAGGAGT 1829
 QY 397 heGlnSerThrLeuSerLysSerGlnValPheThrThrPheLysGlnGluMetGluL 417
 Db 1830 TCAGAACACACTTTCCAAAGCAGGAGGTATTACCAACATTCAGCAGGAGATGGAAA 1889
 QY 417 ySmetThr 419
 Db 1890 AGGTAAT 1897

RESULT 41
 BD060509 529 bp DNA linear PAT 27-AUG-2002
 LOCUS Secreted expressed sequence tags (seSTs).
 DEFINITION
 ACCESSION BD060509
 VERSION BD060509.1 GI:22606115
 KEYWORDS JP 2001518793-A/869.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 529)
 REFERENCE Jacobs,K., McCoy,J.M., Lavallie,E.R., Racie,L.A., Merberg,D.,
 Treacy,M., Spaulding,V. and Agostino,M.J.
 TITLE Secreted expressed sequence tags (seSTs)
 JOURNAL Patent: JP 2001518793-A 869 16-OCT-2001;
 GENETICS INSTITUTE INC
 COMMENT PN JP 2001518793-A/869
 PD 16-OCT-2001
 PF 10-APR-1998 JP 1998543070
 PR 10-APR-1997 US 08/837312
 PI KENNETH JACOBS, JOHN M MCCOY, EDWARD R LAVALLE, LISA A RACIE, PI
 DAVID MERBERG,
 PI MAURICE TREACY, VIKKI SPAULDING, MICHAEL J AGOSTINO PC
 C12N15/12, C12N5/10, C07K14/47, C12Q1/68, A61K38/17 CC Strandedness:
 Double;
 CC Topology: Linear;
 FH Key Location/Qualifiers.
 FEATURES
 source 1..529 Location/Qualifiers
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /db_xref="taxon:4577"
 ORIGIN
 Alignment Scores: 2.18e-30 Length: 529
 Pred. No.: 805.00 Matches: 154
 Score:

Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.35% Mismatches: 0
 Query Match: 28.90% Indels: 0
 DB: 6 Gaps: 0
 US-10-023-523-44 (1-546) x BD060509 (1-529)
 QY 1 MetLysAsnGlnAspLysLysAsnGlnValAlaLysGlnSerAsnProLysSerSerPro 20
 Db 56 ATGAAGACCAACACAAACAAAGACGGGCTGCCAAACATCCATCCAAATCCAAACACGCCCA 115
 QY 21 GlyLnpProGlnuAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
 Db 116 GGACAAACCGGAAGCAGACCCGAGGAGGCCAGGAGCGGCCAGGCGGGCTCTCGA 175
 QY 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
 Db 176 GTGAAGCAGAGGTCCTCCGACAGCAGCCAGGCTCTCCGAAGCCGCGGGGCTCAAGCC 235
 QY 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluLulLeuSerArgGlnLeuGlu 80
 Db 236 AGAACGGCTCAGTCTGGGCGCTTCGTGATGTCCTGAGGAGCTGAGCCGCAACTGGAA 295
 QY 81 AspLileuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAla 100
 Db 296 GACATCTGACACATCTACTGTGTGGACAATAACAGGGGGGGCCCCGGGAGGATGGGCA 355
 QY 101 GlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsn 120
 Db 356 CAGGCTGAGCGGCTGNAACCCGAGATGACAGAGTCCCGGACCTATGTGGCAAGGAT 415
 QY 121 GlyLnpProGluProThrProValValAsnGlyGluLysGluProSerLysGlyAspPro 140
 Db 416 GGGGAGGCTGAACCACTCCAGTAGTCAATGGAGAGAGAACCTCCAGGGGGATCCA 475
 QY 141 AsnThrGluGluLileArgGlnSerAspGluValGlyAspArgAsp 155
 Db 476 AACACAGAGAGATCCGGCAGAGTGACAGGTGGAGACCGAGAA 520
 RESULT 42
 AL671759/c 141790 bp DNA linear ROD 12-JUL-2002
 LOCUS Mouse DNA sequence from clone RP23-151A15 on chromosome 4, complete
 DEFINITION sequence.
 ACCESSION AL671759
 VERSION AL671759.10 GI:22138725
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 141790)
 REFERENCE Direct Submission
 AUTHORS Corby,N.
 TITLE Submitted (12-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
 On Aug 8, 2002 this sequence version replaced gi:21732087.
 COMMENT ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquery@sanger.ac.uk

 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-151A15 is from the RP23-23 Mouse PAC Library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBAC3.6.

FEATURES

Location/Qualifiers
 1..141790
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="4"
 /clone="RP23-151A15"
 /clone_lib="RP23-23"

ORIGIN

Alignment Scores:
 Pred. No.: 2.96e-26 Length: 141790
 Score: 771.00 Matches: 275
 Percent Similarity: 24.49% Conservative: 11
 Best Local Similarity: 23.54% Mismatches: 25
 Query Match: 27.68% Indels: 860
 DB: 10 Gaps: 9

US-10-023-523-44 (1-546) x AL671759 (1-141790)

Qy	233	LysAlaValLeuAlaArgSerLys-----LeuGluSerLeuCysArgGluLeu	248
Db	62348	AAATCGCTGTTGGCAAGACATGGACTTTGGAACTAGTTAGGGGAGCTTTGGCAAGGCTG	62289
Qy	249	GlnArgHisAsnArgSer---LeuLysGluGlyValGlnArgAlaArgGluGluGlu	267
Db	62288	GCAACTATGGACAATCTCTTCTTCAGGAAGAAGCGGTGCAGCGCCCGTGGAGGAG	62229
Qy	268	GluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGln	287
Db	62228	GAGAAGCGCAAGAAGTGAATTCACATTCAGGTGACACTGGAATGACATTCAGCTGCAG	62169
Qy	288	MetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGluLeuAlaGlu	307
Db	62168	ATGGACAGCATACAGCGCAACTCCAGCTGGCCAGGAGATATGGAGCTAGCCGAG	62109
Qy	308	ArgLeuLysLysLeuIleGluGlnTyrGluLeuArgGluGlu-----	321
Db	62108	AGGCTCAAGAAGTTGATCGCAATACGACTTCGTGAGGAGGTGAGCTGCAGGCGATTG	62049
Qy	321	-----	321
Db	62048	AGCCAGCAGCTGGGAGCTGTAGTAGTGCCAGGGGCTGGGGCCGCAACTGCAGGCCT	61989
Qy	321	-----	321
Db	61988	GGGCTTGCTCTTCCTTCAAGAGATGTATCCGTAGATAAGAGCTGAAATAATGGAAAG	61929
Qy	321	-----	321
Db	61928	AAGAGATTACACACACACACACAGTGTATCCCTGGTGTCTGAACCGGACATCTGT	61869
Qy	321	-----	321
Db	61868	AGAGTATTCACTGTAGTAGGAGCGGTGCTCGTCCAGCTCCGCATCAGACCTGTAC	61809
Qy	321	-----	321
Db	61808	ATCAGACCATGCCATTCCAAAGGACATGGGTGGCGAGGAGACCGGATGAAGTGCCTTGT	61749

Qy	321	-----	321
Db	61748	TCTAGAGTTGGGGCTTTCTTGAAGGGAACCTCCTGACTCCATAAATTAACCTTTCTGCA	61689
Qy	322	-HisIleAspLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLysLe	341
Db	61688	GCATATCGACAAAGTCTTCAACATAAAGCACTGCAGCAGCAGCTAGTGTGCGCAAGCT	61629
Qy	341	uGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGlnArgGluLysAspPh	361
Db	61628	CCAGCAGGCCAGGAGATCTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	61569
Qy	361	eLeu-----	362
Db	61568	TGT-GAGGCTCAGGCCCGGCTGTGTGGGCTGGAAGAGTGGGCTGGTCTTGGC	61510
Qy	362	-----	362
Db	61509	TTAGTTCAATCTATTCTTATGATGGAGAGATTTTGCTGCTGCCAATCCAGTAATTGCC	61450
Qy	362	-----	362
Db	61449	TGTCAGACAGCAGCTCTGAGACAGGTAGCTCAGTGGGTCCAGTGTGTGGCTTAAGA	61390
Qy	362	-----	362
Db	61389	AGGTCAGACATACCTCAGGGGCTTCTGACAGGTTCTGGGTCTCTTGGAAATAGCT	61330
Qy	363	-LeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHi	382
Db	61329	CCTGAAGGAAGCGGTGGAGTCCAGAGGATGTGCGAGCTGATGAAGACAGCAGGAGCCA	61270
Qy	382	sLeuLysGlnGlnLeu-----	387
Db	61269	CCTAAAGCAGCAGGT-GAGACGGGAACCTGACCTGCAGTTTCCAGGATGCTTTCACAAG	61211
Qy	387	-----	387
Db	61210	GCCTATCTGAGGGCAGTGGAGACAGGACTGTATAGCCTGGCTTTGTGCCAGCTGGCCT	61151
Qy	387	-----	387
Db	61150	GTTGGTGTGGCCAACTCTTTTATTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	61091
Qy	387	-----	387
Db	61090	CTGTGTAGCCTGGCTGCTCTGAAACTTCTTTGTATACACAGGCTGGCTGGAAACACAGA	61031
Qy	387	-----	387
Db	61030	GATTTCCCTGTCTTAGCCTCTTAAGTAATGGGATTAATGTAGGCATGTGCCATCATGCC	60971
Qy	387	-----	387
Db	60970	AGGCATTTTGTGTCTTTTGAAGACAGGTCTCTGAGTTTAGCTGTGGCTCAACTAAC	60911
Qy	387	-----	387
Db	60910	CCAGAACTTACCTGTAGACAGGCTCTCTGTGGGCTTTGAACCTGTGACATCTCTCTGA	60851
Qy	387	-----	387
Db	60850	CTCCAGATGCATGAGGCGTTAGTACCAGGTCTAGCTCAATTCGTCTACTTTTATTTT	60791
Qy	387	-----	387
Db	60790	GCTGTATTGATATGCATATGAGTGTGTTTTCATGGATGTATATTTTCCCCACTTGTGTT	60731
Qy	387	-----	387
Db	60730	CTGGTACCTTTAGAGTTAGAATAGTTCAGACAGTTGTGGTCTCCATGTAGTGTGGG	60671
Qy	387	-----	387

corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr1>

This sequence is the entire insert of clone RP4-622L5. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP4-622L5 is from the library RPi-4 constructed by the group of Pieter de Jong. For further details see

<http://www.chori.org/bacpac/home.htm>

VECTOR: PCIPAC2.

FEATURES

source

Location/Qualifiers

1..115756
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/map="p34.2-36.11"
/clone="RP4-622L5"
/clone_lib="RPCT-4"

repeat_region
1645..1810
/note="TIGGR1 repeat: matches 2259..2417 of consensus"
repeat_region
2123..2437
/note="TIGGR1 repeat: matches 1957..2259 of consensus"
repeat_region
2726..2782
/note="TIGGR1 repeat: matches 1904..1957 of consensus"
repeat_region
3067..3486
/note="TIGGR1 repeat: matches 1501..1904 of consensus"
repeat_region
3781..4180
/note="TIGGR1 repeat: matches 1117..1501 of consensus"
repeat_region
5053..5082
/note="TIGGR1 repeat: matches 1090..1117 of consensus"
repeat_region
5372..5501
/note="TIGGR1 repeat: matches 966..1090 of consensus"
repeat_region
5628..5704
/note="TIGGR1 repeat: matches 893..966 of consensus"
repeat_region
5991..5492
/note="TIGGR1 repeat: matches 411..893 of consensus"
repeat_region
6800..6835
/note="TIGGR1 repeat: matches 378..411 of consensus"
repeat_region
6968..7022
/note="TIGGR1 repeat: matches 1..378 of consensus"
repeat_region
8281..8358
/note="3 copies 26 mer 79% conserved"
repeat_region
8301..8360
/note="30 copies 2 mer ta 81% conserved"
repeat_region
17674..17777
/note="4 copies 26 mer 77% conserved"
repeat_region
17685..17784
/note="50 copies 2 mer at 82% conserved"
misc_feature
19385..19542
/note="match: GSS: Em:AQ306911"
repeat_region
19551..19574
/note="12 copies 2 mer tt 100% conserved"
repeat_region
20865..20922
/note="29 copies 2 mer at 70% conserved"
repeat_region
20934..21043
/note="55 copies 2 mer at 70% conserved"
repeat_region
20940..21043
/note="4 copies 26 mer 74% conserved"
repeat_region
21168..21234

/note="tRNA-asp-GAC repeat: matches 1..75 of consensus"
28263..28691
/note="match: GSS: Em:AQ317905"
complement(29964..30447)
/note="match: GSS: Em:AQ732256"
30465..30718
/note="match: STS: Em:GS8909"
match: GSS: Em:AQ376964"
38923..39446
/note="L1MEC repeat: matches 1546..2095 of consensus"
39761..39882
/note="L1MEC repeat: matches 1429..1546 of consensus"
complement(43254..43726)
/note="match: GSS: Em:AQ146371"
45555..46050
/note="MER44C repeat: matches 1..728 of consensus"
46055..46468
/note="match: GSS: Em:AQ040362"
48914..48982
/note="MER321B repeat: matches 718..790 of consensus"
48985..49173
/note="MER21B repeat: matches 77..263 of consensus"
49285..49651
/note="match: GSS: Em:AQ761532"
52462..53177
/note="match: GSS: Em:AQ040082"
54716..54865
/note="match: GSS: Em:AQ680771"
60551..60614
/note="32 copies 2 mer tt 70% conserved"
complement(62219..62382)
/note="match: STS: Em:G20585"
complement(66030..66267)
/note="match: GSS: Em:AQ514287"
complement(66041..66267)
/note="match: GSS: Em:AQ803288"
66066..66255
/note="match: GSS: Em:B69239"
68882..69541
/note="CpG island"
/evidence=not_experimental
71427..71448
/note="11 copies 2 mer tt 100% conserved"
73527..74249
/note="match: GSS: Em:AQ892680"
74514..74926
/note="match: GSS: Em:AQ441232"
74532..74991
/note="match: GSS: Em:AQ812740"
74626..75008
/note="match: GSS: Em:AQ091614"
complement(74682..75020)
/note="match: STS: Em:L29744"
complement(74683..74951)
/note="match: STS: Em:L29729 Em:L29763"
74717..74867
/note="match: STS: Em:G15744"
82692..83089
/note="match: STS: Em:G09447"
83391..83481
/note="match: GSS: Em:AZ310334"
85284..85337
/note="27 copies 2 mer tt 88% conserved"
85339..85718
/note="match: STS: Em:G28385"
complement(87250..87612)
/note="match: STS: Em:G06968"
complement(87455..87579)
/note="match: STS: Em:G26511"
complement(87851..88313)
/note="match: GSS: Em:AQ536603"
89654..89685
/note="16 copies 2 mer tc 87% conserved"

Db 2100 ATGAAGAACCAAGACAAAAGAACCGGGCTGCCAAACAATCCAAATCCAAAAGACGCCCA 2159
 QY 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
 Db 2160 GGACACCGGACAGGACCCGAGGAGCCAGGAGCGCCAGCCAGCGGCTCTGCA 2219
 QY 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgGlyProGlu 56
 Db 2220 GTAGAACGACGAGGTCCCGCAGCAGCCAGGCTCTCTCGGAAGCCGAGGGTGTGCCAG 2279
 QY 56 56
 Db 2280 CTCTGCGTTCCACGCGGAGGGGAGGAGCTGTGGGTCGGCTCGCTTCTGGACTTAC 2339
 QY 56 56
 Db 2340 AGCCGAGGCCAGGTTGTCCGGAGGAGGAGATGTAGAATGAGAGACAGTGTCTGGGGC 2399
 QY 56 56
 Db 2400 CGCGGTCCCGCTGCGCTCTGGCGAGTGTGGGAGCTGCCCTCTAAGCACAGGAACAG 2459
 QY 56 56
 Db 2460 AGTTCTGGAGAGAGCTCCGACGGGATTAAGTCAGGTGGCAGCAACAGGACCCAGT 2519
 QY 56 56
 Db 2520 CAGGAATCCAGGTCCTCGTTAGAAACACCTCAGCCACAGCAGCTAAGTGCCTTCTGT 2579
 QY 56 56
 Db 2580 TTCAGGCATTTCTAGATGATCTGATGCGAAGATGGTTTGTGGGGGAGGAGAGA 2639
 QY 56 56
 Db 2640 TGGACTAGAAGTTGCTCCGTGCCATCCCTGTGTGTGATGCTTTATCATATTTATGATC 2699
 QY 56 56
 Db 2700 TAACAAATATGTTCCGGTGTAGTAGAATAGTTGTGTCTTTACAGTAAACAGACT 2759
 QY 56 56
 Db 2760 TAAAGAAGTTAGCAACGATTACTATAATTCTTGATTTTAAAGATGTTTCAATCTAAA 2819
 QY 56 56
 Db 2820 TTCTGACAGAACTAGATTGTCGAATGATCTCCATCTTCTGCTCTCTCATTTCCATAAA 2879
 QY 56 56
 Db 2880 AAAAAAGTTAGCAACATTTAACTCAACTGATGATTTGGCTGGCCCTGAAAAATCCC 2939
 QY 57 57
 Db 2940 AACCACTGATATAATCGCTCTTCTTCTACTCTACCCCTCATCTCTCTCTGTAGGGG 2999
 QY 58 AlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArg 77
 Db 3000 CT-CAGCCAGAACGCTCAGTCTGGGGCCCTTCGTGATGCTCTGAGGAGCTGAGCCGC 3058
 QY 78 GlnLeuGluAspIleLeuSerThrTyrcysValAspAsnAsnGlnGlyProGlyGlu 97
 Db 3059 CAACCTGAAGACATCTAGACACATCTGTGTGGACAATAACACGGGGCGCCCGCCGAG 3118
 QY 98 AspGlyValGlnGlyGluProAlaGluProGluAspAlaGluLeuSerArgThrTyrcysVal 117
 Db 3119 GATGGGGCAGAGGTGAGCGCTGAGCCCGAGATGCGAGAGAGTCCCGACCTATGTG 3178
 QY 118 AlaArgAsnGlyGluProGluProThrProValValAsnGlyGluLeuProSerLys 137
 Db 3179 GCAAGGAATGGGAGCTCGAACCACTCCAGTAGTCAATGGAGAGAGGAACCCCTCCAG 3238

QY 138 GlyAspProSerThrGluGluLeuLeuArgGlnSerAspGluValGlyAspArgHisArg 157
 Db 3239 GGGGATCCAAACACAGAGAGATCCCGCAGAGTCCAGGTCGAGAGCCGAGACCATCGA 3298
 QY 158 ArgProGlnGluLysLysAlaLysGlyLeuGlyLysGluLeuThrLeuLeuMetGln 177
 Db 3299 AGGCCACAG 3358
 QY 178 ThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuLeuCysLysLysTyrcysAla 197
 Db 3359 3359
 QY 198 GluLeuLeu 3554 bp mRNA linear PRI 30-SEP-2003
 Db 3401 AGTCTGTCTTCCAGGATTCAGAGAAACGGTACTTCTCAGAGACAGC---AAGTCACCTC 3457
 QY 214 GlnLysLysGlnSerGln 219
 Db 3458 TAGTCTAATCAAGCCAG 3475
 RESULT 46
 BC046565
 LOCUS
 DEFINITION Homo sapiens taxilin, mRNA (cdna clone IMAGE:5534975), complete cds.
 ACCESSION BC046565
 VERSION BC046565.1 GI:28302155
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 3554)
 AUTHORS Klausner R.D., Collins, P.S., Wagner, L.H., Derge, J.G., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L., Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Frange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McSwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.C., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 MEDLINE 22388257
 PUBMED 12477932
 REFERENCE 2 (bases 1 to 3554)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (10-FEB-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 REVARK NIH-MGC Project URL: http://mgc.nci.nih.gov
 COMMENT Contact: MGC help desk
 Email: gcgaps-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural Gaithersburg, Maryland;

4	GlnAspIysIysAsnGlyAlaAlaLysGlnSerAsnProLysSerSerProGlyGlnPro	23
49	CAACCAAAAGATGAAATGAAATGACAGATGTTCAACTCCGCTAAGTCCACAAATGAG	108
24	GluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAlaValGluAla	43
109	GAACCGCTCCACTCAGTGTTCGAACGGCCACCAACACGAGAGATGTGACCAATATT	168
44	GluglyProGlySerSerGlnAlaProAaGlySProGluGlyAlaGlnAlaArgThrAla	63
169	TCATTCT--TCGTCATGCAGCTCTCCTACGACACCT--	201
64	GlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGluAspIleLeu	83
202	-----AGGCATGTCATCA-----TTG	219
84	SerThrTyrCysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAlaGlnGlyGlu	103
220	GCAGGGTAC-----TCAGCTCTGCC-----TCAGCTCTGCC	240

QY 296 SerLysLeuArgGlnGlnAsnVetGluLeuAlaGluArgLeuLysLeuIleGln 315
 DB 466 ATCAAGCTCGCGATTACAACTCAGATGACGACCAAGAGTTAAAGCTCTGCGCCGAG 525
 QY 316 TyrGluLeuArgGlnGlnHisIleAspLysValPheLysHisLysAspLeuGlnGln 335
 DB 526 TACCAACAGGAGGAGCAGACCTCGGAGAGCTCAATGAACAGGTTCAAGTTGGAGGCCAG 585
 QY 336 LeuValAspAlaLysLeuGlnGlnAlaGlnGlnMetLeuLysGluAlaGluLysHis 355
 DB 586 CTGCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 645
 QY 356 GlnAsgLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeu 375
 DB 646 ACAGAGAAACACAGATTGGCTTGAGAACTGATGCAAGCGCACGTCGCATAAGGAT 705
 QY 376 MetLysGlnGlnGlnThrHisLysGlnGlnLeuAlaLeuTyrThrGluLysPheGlu 395
 DB 706 CTCACCGATCGGAGCAGCACTCAAGGAGCACTGAACATATACACCGCAAGTACGAT 765
 QY 396 GluPheGlnAsnThrLeuSerLysSerGluValPheThrThrPheLysGlnGlnMet 415
 DB 766 GACTTCACAGCTGCTGCGAGAGTCAATAGGTTTCGCGAGCTACAGGTGGAGCTT 825
 QY 416 GluLysMetThrLysLysLysLysLysLysGluLysGluThrThrMetTyrArgSerArg 435
 DB 826 GAGAAGATGTCACAGCACACTAAGAGATCGAGAGGAGGCGCTGGCTGGCGGCAAAAG 885
 QY 436 TrpGluSerSerAsnLysAlaLeuGluMetAlaGluLysThrValArgAspLys 455
 DB 886 TAGCAGAAAGCTAATCCCTGATGTCGATTAGCCACAGAAAGAGTCTGCAGCGGAG 945
 QY 456 GluLeuGluGluGlnValLysLysLysLysLysLysGluLysLysCysArgAlaLeuGln 475
 DB 946 CACTCAGAGCGCTGACAGCAATACAGAGCTGCAGAGCTACTGGTGCTCTACAG 1005
 QY 476 ThrGluArgAspLysLysLysLysLysLysValGlnApp 487
 DB 1006 CTGAGAGGACCACTGCACAGTCTGCTGGCGGAT 1041
 RESULT 50
 AC111774
 LOCUS
 DEFINITION Rattus norvegicus clone CH230-171G2, WORKING DRAFT SEQUENCE, 10
 unordered pieces.
 AC111774
 ACCESSION
 VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 KEYWORDS
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 245095)
 Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, P.,
 Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buha, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
 Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,

Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogue, M.,
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorensheva, L., Louised, H., Lozano, R.J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindartine, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mahoney, S., McLeod, M.P., McNeill, T.Z., Meenan, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwackeleneh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Shen, H.,
 Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Steidle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
 Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstein, G., and Gibbs, R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 245095)
 Worley, K.C.
 Direct Submission
 Submitted (19-FEB-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 245095)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (09-NOV-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Nov 9, 2002 this sequence version replaced gi:23321701.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GOBO
 Center clone name: CH230-171G2
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 153773 bases at least Q40
 Consensus quality: 160170 bases at least Q30
 Consensus quality: 164659 bases at least Q20
 Estimated insert size: 160182; sum-of-contigs estimation
 Quality coverage: 10x in Q20 bases; sum-of-contigs estimation

Job time : 6777.85 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 8, 2004, 06:37:58 ; Search time 3984.89 Seconds
(without alignments)
4091.647 Million cell updates/sec

Title: US-10-023-523-44

Perfect score: 2785

Sequence: 1 MKNQDKKNGAKQSNPKSSP.....APSTRASGQTGPQETSARA 546

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 50 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10023523/runat_08062004_063749_29950/app.query.fasta_1.1422
-DB=EST -Qfmt=FASTAP -SUFFIX=rst -MINMATCH=0.1 -ICOPCL=0 -LOOPEXT=0
-UNITS=bites -START=1 -END=1 -MATRIX=blosun62 -TRANS-human40.cdi -LIST=50
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTFMT=pcr -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10023523@cgn_1_15307@runat_08062004_063749_29950 -NCPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*

1: em_estba:*
2: em_estba:*
3: em_estin:*
4: em_estin:*
5: em_estov:*
6: em_estov:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estin:*
16: em_estin:*
17: em_gss_hum:*
18: em_gss_hum:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_pro:*
26: em_gss_rtd:*
27: em_gss_ohg:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1295	46.5	817	12	BG761234	BG761234 602718583
2	1294	46.5	2797	11	AK031783	AK031783 Mus muscu
3	1294	46.5	2832	11	AK044130	AK044130 Mus muscu
4	1294	46.5	3810	11	AK030100	AK030100 Mus muscu
5	1279	45.9	798	14	CA328534	CA328534 UI-M-FYO-
6	1269	45.6	775	14	CF728426	CF728426 UI-M-HB0-
7	1264	45.4	770	14	CF735205	CF735205 UI-M-HB0-
8	1251.5	44.9	806	14	CA320886	CA320886 UI-M-FW0-
9	1250.5	44.9	932	12	BG827157	BG827157 602751143
10	1243.5	44.6	903	13	EX752507	EX752507 BX752507
11	1242.5	44.6	779	14	CB247992	CB247992 UI-M-FYO-
12	1234	44.3	942	13	BQ714219	BQ714219 AGENCOURT
13	1233	44.3	945	9	AL961454	AL961454 AL961454
14	1209	43.4	725	13	BQ701958	BQ701958 UI-M-FYO-
15	1207	43.3	735	13	BQ769910	BQ769910 UI-M-FYO-
16	1205	43.3	759	14	CF995807	CF995807 AGENCOURT
17	1192	42.8	1608	11	AK084639	AK084639 Mus muscu
18	1191	42.8	2480	11	AK085904	AK085904 Mus muscu
19	1170	42.0	799	13	BQ849604	BQ849604 AGENCOURT
20	1169.5	42.0	750	14	CA327097	CA327097 UI-M-FYO-
21	1162	41.7	727	13	BQ770181	BQ770181 UI-M-FYO-
22	1153.5	41.4	716	14	CF728006	CF728006 UI-M-HB0-
23	1148.5	41.2	819	10	BF783468	BF783468 60211238
24	1136.5	40.8	4452	11	AK031023	AK031023 Mus muscu
25	1136	40.8	778	14	CA324134	CA324134 UI-M-FYO-
26	1135.5	40.8	727	14	CB244426	CB244426 UI-M-FYO-
27	1116	40.1	699	13	BQ701604	BQ701604 UI-M-FYO-
28	1099	39.5	780	13	EX855901	EX855901 BX855901
29	1099	39.5	930	13	BQ950691	BQ950691 AGENCOURT
30	1081	38.8	702	9	AL637771	AL637771 AL637771
31	1071	38.5	735	12	BM963869	BM963869 UI-M-EQ0-
32	1070	38.4	698	12	BQ618489	BQ618489 BJ618489
33	1069.5	38.4	1318	12	BM455349	BM455349 AGENCOURT
34	1061	38.1	861	12	B1102887	B1102887 602888485
35	1060	38.1	647	14	CD578468	CD578468 UI-M-FYO-
36	1059	38.0	653	13	BU613886	BU613886 UI-M-EW0-
37	1052.5	37.8	722	14	CA510893	CA510893 UI-R-FYO-
38	1046	37.6	784	14	CB247697	CB247697 UI-M-FYO-
39	1041.5	37.4	796	12	BQ975353	BQ975353 602842995
40	1040	37.3	665	9	AL887857	AL887857 AL887857
41	1037	37.2	782	13	BQ231392	BQ231392 603947670
42	1037	37.2	862	14	CA793605	CA793605 AGENCOURT
43	1034	37.1	654	13	BQ389733	BQ389733 NISC mq09
44	1014.5	36.4	954	13	BU151934	BU151934 AGENCOURT
45	1005.5	36.1	777	14	CA247310	CA247310 UI-M-FYO-
46	1005.5	36.1	780	14	CA328385	CA328385 UI-M-FYO-
47	993.5	35.7	890	12	B1661586	B1661586 603305711
48	991.5	35.6	965	10	BF796069	BF796069 602258906
49	990	35.5	686	13	BU054710	BU054710 UI-M-FD0-
50	988	35.5	1611	29	AY405165	AY405165 Homo sapi

ALIGNMENTS

RESULT 1
BG761234
LOCUS
DEFINITION BG761234 602718583F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4858388 5',
mRNA sequence.
ACCESSION BG761234
VERSION BG761234.1 GI:14071887
KEYWORDS EST,
EST

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 817)
 AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC/DCTD/DTF
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLC1712 row: f column: 21
 High quality sequence stop: 766.
 FEATURES
 Location/Qualifiers
 1..817
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:485838"
 /tissue_type="melanotic melanoma, high MDR (cell line)"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 49"
 /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GCCACGAG(G). Size-selected >500bp for average insert size
 1.8kb. Library constructed by Ling Hong in the laboratory
 of Gerald M. Rubin (University of California, Berkeley)
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC
 Library."

ORIGIN
 Alignment Scores:
 Pred. No.: 2,02e-73 Length: 817
 Score: 1295.00 Matches: 263
 Percent Similarity: 97.41% Conservative: 0
 Best Local Similarity: 97.41% Mismatches: 3
 Query Match: 46.50% Indels: 4
 DB: 12 Gaps: 0
 US-10-023-523-44 (1-546) x BG761234 (1-817)

QY 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerSerPro 20
 Db 15 ATGAAGAACCAACACAAAGACCGGGCTGCCAAACATCCATCCAAAGACGCCCA 74
 QY 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
 Db 75 GGACACCGGACGACGAGACCCCGAGGGAGCCCGACGCGCCAGCGGCTCTGCA 134
 QY 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
 Db 135 GTAGAGACGAGAGGTCTCCCGACGACGACCGAGCTCTCGGAGCGCGGGGCTCAAGCC 194
 QY 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
 Db 195 AGAACGGCTCAGTCTGGGGCCCTTCGTGATGCTCTGAGGAGCTGAGCGCCCACTGGAA 254
 QY 81 AspIleLeuSerThrTyrcysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAla 100
 Db 255 GACATCTAGCACAATCTGTGTGGACAATAACACGAGGGGGCCCGCGGAGGATGGGCA 314
 QY 101 GlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrcysValAlaAlaArgAsn 120
 Db 315 CAGGGTGAGCGCGGTGAACCCGAGAGATGCACAGAGTCCCGGACCTATGTGGCAAGGANT 374

QY 121 GlyGluProGluProThrProValValAsnGlyGluLysGluProSerLysGlyAspPro 140
 Db 375 GGGGAGCGCTCAACCACTCCAGTAGTCAATGGAGAGAGAACCCCTCCAAAGGGGGATCCA 434
 QY 141 AsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAspHisArgArgProGln 160
 Db 435 AACACAGAAAGATCCGGCAGAGTCCAGAGTCCGAGACCGGAGACCATCGAAGGCCACAG 494
 QY 161 GluLysLysLysAlaLysGlyLeuGlyLysGluLysGluLeuMetGlnThrLeuAsn 180
 Db 495 GAGAGAAAAGCCAGGGGTTGGGAGGAGATCACGTTGCTGATGCAGACATTGAAT 554
 QY 181 ThrLeuSerThrProGluGluLysLeuAlaAlaLeuLysLysLysLysLysLysLys 200
 Db 555 ACTCTGAGTACCCAGAGGAGAGTGGTCTCTGTGCAAGAAAGTATGCTGAAGTCTG 614
 QY 201 GluGluHisArgAsnSerGlnLysGlnMetLysLeuLysGlnLysLysLysLysLys 220
 Db 615 GAGGAGCACCGGATTCACAGAGCAGATGAAGCTCTACAGAAAAGCAGAGCCAGCTG 674
 QY 221 -ValGlnGluLys-AspHisLeuArgGlyGluHisSerLysAlaVal-LeuAlaArg- 239
 Db 675 GGTCAAGAGAGAGGACCACTGCGCGGTGAGCAGACAGAGGCGCTCTTGGCCCGCCAG 734
 QY 239 rLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGlu 259
 Db 735 CAAGCTTGAGAGTCTATGCTCCGTGAGTGCAGCGGCACAAACGATCCTCAAGAGAGG 794
 QY 259 yValGlnArgAlaArgGluGlu 266
 Db 795 TGTCCAGCGGGCCCGCGAGGAA 816

RESULT 2
 AK031783 2797 bp mRNA linear HTC 18-SEP-2003
 LOCUS Mus musculus 11 days embryo head cDNA, RIKEN full-length enriched
 DEFINITION library, clone:6230424F16 product:hypothetical protein, full insert
 sequence.
 ACCESSION AK031783
 VERSION AK031783.1 GI:26327606
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE 2
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159
 REFERENCE 3
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
 Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861
 REFERENCE 4


```

Db      878 GCATTAAGGAGACCATATTGATTAAGTATTCAACACAGGAATTCACACACAGCTT 937
Qy      337 ValAspAlaLysLeuGlnGlnAlaGlnMetLeuLysGluAlaGluGluGluGlnHisGln 356
Db      938 GTGGATGCCAACTTCAGCAACACACAGCTGATAAAGAGAGCTGATGAATAACATCAG 997
Qy      357 ArgGluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMet 376
Db      998 AGAGAGAGAGAGCTTTTATTAAAGAGACACAGGAATCCAGGCAAAATATGAAGAAATG 1057
Qy      377 LysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuLysThrGluLysPheGluGlu 396
Db      1058 AAACAGCAGAGAGTCAACTAAACACAGAGCTTCTTTATATGATAATTTGAAGAA 1117
Qy      397 PheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGlu 416
Db      1118 TTCCAGACTACTATGCAAAAGCAATGAACCTTTTACAACCTTCAGGCAGGAATAATGAA 1177
Qy      417 LysMetThrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 436
Db      1178 AAGATGACAAAGAAATTAATAAACTGGAAGAAAGAAACAAATAATATGCGGTACCAATGG 1237
Qy      437 GluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGlu 456
Db      1238 GAAACAATAATAAGCACTTCTGCAGATGCGCGAAGAGAAACCTGTCGTGATAAAGAG 1297
Qy      457 LeuGluGluLeuGlnValLysLysLysLysLysLysLysLysLysLysLysLysLysLys 476
Db      1298 TACAAGGCTTTTCAATAAACTGGAAGCGTTAGAGAGAGCTGTGCAGGCTCTTCAGACA 1357
Qy      477 GluArgAsnAspLeuAsnLysArgValGlnAspLeu 488
Db      1358 GAGAGAAATGAGCTCAACAGAGAGTCAAGTCTG 1393

RESULT 3
AK044130
LOCUS
DEFINITION
Mus musculus 10 days neonate cortex cDNA, RIKEN full-length
enriched library, clone:A830093C13 product:hypothetical protein,
full insert sequence.
AK044130
VERSION
AK044130.1 GI:26336211
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carrinci, P. and Hayashizaki, Y.
2
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
9279253
MEDLINE
10349636
PUBMED
3
Carrinci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Iton, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
JOURNAL
11042159
MEDLINE
4
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kusunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujitake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, T., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system -384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)

```

MEDLINE PUBMED REFERENCE AUTHORS

TITLE JOURNAL REFERENCE AUTHORS

TITLE JOURNAL REFERENCE AUTHORS

TITLE JOURNAL

COMMENT

FEATURES source

CDS

ORIGIN

```

20530913
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690. (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6
(bases 1 to 2832)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, T., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takanashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.
Location/Qualifiers
1. 2832
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:A830093C13"
/db_xref="MGI:2409425"
/db_xref="taxon:10090"
/clone="A830093C13"
/tissue_type="cortex"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 days neonate"
64. 1638
/notes="unnamed protein product; hypothetical protein
(evidence: rscDS, ProCrest, decoder, NCBI CDS
Predictor, Longest-ORF)
putative"
/codon_start=1
/protein_id="BAC31791.1"
/db_xref="GI:26336212"
/translators="MATRILEVTRGRGGTTEASGEGRRRRSPQKFEIGTMEPAR
ICGLGKADMCVQNDILHODPSCGTTKTKHSLGEGSDFTIKRNILVSVFCT
QKREPIPREARTGPDGQDQSECSKEKTKGKVLMLQALNTLSTPEKLAALC
KKTADLESNRVQKQKILQKQQAQIVKVKHLOSEHSAKILAPSKLSLRELQRL
NKTLEENMQQAREEERREKTAHFQIILNEIQAGLEQDTHNKLRENELGKRL
KXLEIYALREEDHDKFKHKLQQLQVDAKQQTOLIKADEKQREKRELFKRL
ESHKYEQKQKQVQLKQSLYDFEFTMAKSNELFTTFQEMSKRTTKIKKL
EKETITWRTKNNKALLQMAEKTRVDRKVFQIKLERLEKLCRALQTERNELNE
KSVLKEQVSIKAADGDLVSPAQCAVLDSFKTSRTTLGMHLEARKSVCEKSAQ
KPSSSGSPAQIESVD"

```



```

791 TTCCAGATAACTCTAAATGAAATCCAAAGCTCAGTTGGAAACACATGACATCCACAATGCC 850
QY 297 LysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLeuLysLeuLysLeuLys 316
Db 851 AAATCTGGCAGCAGGAGAAATTAAGTGGAGGAGAGTTGAGAAAGCTTATTGAGCAGTAT 910
QY 317 GluLeuArgGlnGluHisLeuLysValPheLysHisLysAspLeuGlnGlnLeu 336
Db 911 GCACCTAAGCGAAGAGCATATTGATTAAGATTATCAACACACAGGAATTCACACACAGCTT 970
QY 337 ValAspAlaLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 356
Db 971 GTGGATGCCAACTTCAGCAACACACAGCTGTATAAAGAGAGCTGATGAATAAATCATCAG 1030
QY 357 ArgGluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMet 376
Db 1031 AGAGAGAGAGACTTTTATTAAAGAGACACAGATCCAGGCACAAATATGACAAATG 1090
QY 377 LysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 396
Db 1091 AAACAGCAAGAGTAGTACAACTAAACACAGCAGCTTCTCTTTATATGATAAATTTGAAGA 1150
QY 397 PheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGlnMetGlu 416
Db 1151 TTCCAGACTACTATGCGAATAAGCAATGAACTTTTACACCTTCAGGCAGAAATGGAA 1210
QY 417 LysMetThrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 436
Db 1211 AAGATGACAAAGAAATTAATAAAGTGGAAAGAAACAATAATATGCGTACCAATATG 1270
QY 437 GluSerSerAsnLysAlaLeuLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 456
Db 1271 GAAACAAATAATAAGACACTCTCGAGATGGCCGAGAGAAACACTGCTCGTAAAGAG 1330
QY 457 LeuGluGlyLeuGlnValLysLysLysLysLysLysLysLysLysLysLysLysLys 476
Db 1331 TACAAGGCTTTCAATAAATAAGTGAACGGTTAGAGAGCTGTGACGGCTTTCAGACA 1390
QY 477 GluArgAsnAspLeuAsnLysArgValGlnAspLeu 488
Db 1391 GAGAGAAATGAGCTCAACGAGAAGTGAAGTCTTG 1426

RESULT 5
CA328534 798 bp mRNA linear EST 09-JUL-2003
LOCUS UI-M-F10-cex-h-10-0-UI.r1 NIH_BMAP_F10 Mus musculus cDNA clone
DEFINITION IMAGE:6825755 5', mRNA sequence.
ACCESSION CA328534
VERSION CA328534.1 GI:24546632
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 798)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
Location/Qualifiers
1..798
/organism="Mus musculus"

```

```

/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6825755"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_F10"
/notes="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:751-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCAGACAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

```

ORIGIN

```

Alignment Scores:
Pred. No.: 2,08e-72 Length: 798
Score: 1279.00 Matches: 258
Percent Similarity: 98.86% Conservative: 2
Best Local Similarity: 98.10% Mismatches: 3
Query Match: 45.92% Indels: 1
DB: 14 Gaps: 0

US-10-023-523-44 (1-546) x CA328534 (1-798)

QY 136 SerLysGlyAspProAsnThrGluGluLeuArgLysSerAspGluValGlyAspArgAsp 155
Db 10 TCTAAGGAGAGAGCTTGAACACAGAGAGATCCGAGCAGTGATGAGTTGAGACCCAGAC 69
QY 156 HisArgArgProGlnGluLysLysLysAlaLysGlyLeuGlyLysGluLeuThrLeuLeu 175
Db 70 CATCGGAGGCCACAGGAGAGAGAAAGCAAGGGCTAGGGAAGGAGATCACTCTGCTG 129
QY 176 MetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuLysLysLys 195
Db 130 ATGCAGACACTGAACAGCTGAGTACCCAGAGAGAGAGCTGCTGCTGCTGCAAGAAG 189
QY 196 TyrAlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLys 215
Db 190 TATGCTGAGCTGCTGGAAGAGCATCGGAACCTCCAGAGCAGATGAAGCTCTGCGAGAG 249
QY 216 LysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaVal 235
Db 250 AAGCAGAGCCAGCTCGTCGAGGAGAGAGGACCATCTGCGAGGGGGAACACACAGAGGCTGC 309
QY 236 LeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeu 255
Db 310 CTGGCCGAGAGCAGCTTGGAGTCTGTGCGGAGAGCTGCAACGGCACACACCGTCCCTG 369
QY 256 LysGluGlyValGlnArgAlaArgGluGluGluGluLysArgLysGluValThrSer 275
Db 370 AAGGAGAGAGCGCTGAGCGAGCGCCGCTGAGGAGAGAGAGAGGAGGAGGAGAGGAGTCA 429
QY 276 HisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsn 295
Db 430 CATTCCAGGTGACACTGAATGATCATCTAGCTGTCAGATGGAACAGCATACGAGGAGAAAC 489
QY 296 SerLysLeuArgGlnGlnAsnMetGluLeuAlaGluArgLeuLysLysLeuLeuGln 315
Db 490 TCCAAGCTGCGCCAGAGAGATATGGAGCTAGCCGAGAGGCTCAAGAAGTTGATCGAGCAA 549
QY 316 TyrGluLeuArgGlnGluHisLysLeuValPheLysHisLysAspLeuGlnGlnGln 335

```


REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 770)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. James Lin University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: PYX-5

FEATURES

source

Location/Qualifiers
 1..770
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:30615737"
 /tissue_type="whole eye"
 /dev_stage="embryo 12.5,13.5,14.5 dpc"
 /lab_host="DH10B (TI phage resistant)"
 /clone_lib="NIH BMAP HBO"

/note="Organ: Eye; Vector: pYX-Asc; Site 1: Ecor I;
 Site 2: Not I; The library was constructed according
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with Ecor I adaptor, digested with NotI and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is TATTGAGT. This library was created for the University
 Iowa Brain Anatomy Project (BMAP). Gene Discovery in the
 Developing Mouse Nervous System", supported by National
 Institute of Mental Health (NIMH)."

ORIGIN

Alignment Scores:

Pred. No.: 1,83e-71 Length: 770
 Score: 1264.00 Matches: 253
 Percent Similarity: 99.22% Conservatives: 1
 Best Local Similarity: 98.83% Mismatches: 2
 Query Match: 45.39% Indels: 0
 DB: 14 Gaps: 0

US-10-023-523-44 (1-546) x CF735205 (1-770)

QY 163 LysLysAlaLysGlyLeuGlyGlyGlyLeuMetGlnThrLeuAanThrLeu 182
 Db 2 AAGAAGCCAGGGGCTAGGAGAGATCACTCTGCTGATGAGACACTGAACAGCTG 61
 QY 183 SerThrProGluGluLeuAlaLeuCysLysLysTyrAlaGluLeuGluGlu 202
 Db 62 AGTACCCAGAGGAGAGCTGGCTGCTGCAGAGATGCTGAGCTGCTGGAAGAG 121
 QY 203 HisArgAsnSerGlnLysGlnMetLysLeuGlnLysLysGlnSerGlnLeuValGln 222
 Db 122 CATCGGAATCGCAGAGACATGATGAGCTCTGAGAGAGAGAGAGAGAGAGAGAG 181
 QY 223 GluLysAspHisLeuArgGlyGlyHisSerLysAlaValLeuAlaArgSerLysLeuGlu 242
 Db 182 GAGAGAGACCATCTCGAGGGGACACAGCAGAGCTGTCTGGCCGAGCAAGCTTGAG 241
 QY 243 SerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGluGlyValGlnArg 262
 Db 242 AGTCGTCCCGAGCTCAACGGCAACCGGTCTTGAAGAGAGAGAGAGAGAGAGAG 301
 QY 263 AlaArgGluGluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsn 282

Db 302 GCCCGTAGGAGGAGGAGCGCAAGAGAGTGAATTCACACTTCCAGGTGACACTGAAT 361
 QY 283 AspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsn 302
 Db 362 GACATTGAGCTGAGATGGACACATACGAGCGAACTCCAGCTGCCCGAGGAGAT 421
 QY 303 MetGluLeuAlaGluArgLeuLysLeuIleGluGlnTyrGluLeuArgGluGluHis 322
 Db 422 ATGAGAGCTAGCGGAGAGCTCAAGAGTTGATCGAGCAATACGAGCTTCGTGAGGAGCAT 481
 QY 323 IleAspLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGln 342
 Db 482 ATCGACAAAGCTTCAACATAGGACCTCGAGCAGCAGGCTAGTGGACGCCAGCTCCAG 541
 QY 343 GlnAlaGlnGluMetLeuLysGluAlaGlnGluArgHisGlnArgGluLysAspPheLeu 362
 Db 542 CAGGCCAGGAGATGCTGAAGAGGAGAGAGCGGCCAGGCCAGGAGAGAGAGAGAGAG 601
 QY 363 LeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHis 382
 Db 602 CTGAGGAGAGCGGTGGAGTCCAGAGGATGTGGAGCTGATGAGCAGCAGGAGAGAGCCAC 661
 QY 383 LeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSer 402
 Db 662 CTAAAGCAGCAGCTCGGCTGTACACGAGAGAGTTTGGAGAGTTCCAGACACACTTCC 721
 QY 403 LysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMet 418
 Db 722 ANAGCAGTGAAGTGTCCACCGCTTCAACAGGAGATGGNAAAGATG 769

RESULT 8

CA320886

LOCUS

DEFINITION

IMAGE:6817313 5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CA320886 806 bp mRNA linear EST 09-JUL-2003
 UI-M-FW0-ccb-h-16-0-UI.r1 NIH BMAP_FW0 Mus musculus cDNA clone
 IMAGE:6817313 5', mRNA sequence.

CA320886

CA320886.1 GI:24538984

EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 806)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)

Seq primer: PYX-5

FEATURES

source

Location/Qualifiers
 1..806
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6817313"
 /tissue_type="whole brain"
 /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
 /lab_host="DH10B (TI phage resistant)"
 /clone_lib="NIH BMAP_FW0"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: Ecor I;
 Site 2: Not I; The library was constructed according
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I and then cloned directionally into pX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is "AGCAGACGAC". This library was created for the University of Iowa Brain Anatomy Project (BMAP): "Gene Discovery in the Developing Mouse Nervous System", supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:
Pred. No.: 1-22e-70 Length: 806
Score: 1251.50 Matches: 259
Percent Similarity: 97.39% Conservative: 2
Best Local Similarity: 96.64% Mismatches: 6
Query Match: 44.94% Indels: 3
DB: 14 Gaps: 1

US-10-023-523-44 (1-546) x CA320886 (1-806)

Qy 123 ProGluPro---ThrProValValaengGlyGluLysGluProSerLysGlyAspProAsn 141
Db 3 CCTGAACCGAGGATTCAGTCGTCAACGCGAGAGGAGACCTCTAAGGAGAGCCCTGGA 62
Qy 142 ThrGluGluileArginSerAspGluValGlyAspArgAspHisArgArgProGluGlu 161
Db 63 ACAGAGGAGATCCGAGCGAGTGATGATGAGTGGAGACCCAGACCATCGAGGCCACAGAG 122
Qy 162 LysLysLysAlaLysGlyLysGlyLysGluLysLysLysLysLysLysLysLysLysLys 181
Db 123 AAGAGAGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 182
Qy 182 LeuSerThrProGluGluLysLysLysLysLysLysLysLysLysLysLysLysLysLys 201
Db 183 CTGAGTACCCAGAGAGAGAGTGGCTGCTGCTGCAAGAGTATGCTGAGTCTGCTGAA 242
Qy 202 GluHisArgSerGlnLysGlnMetLysLysLysLysLysLysLysLysLysLysLysLys 221
Db 243 GAGCATCGGAATCCGAGAGAGAGATGAGTAACTCTCTGCAAGAGAGAGAGAGAGAGAG 302
Qy 222 GlnLysLysAspHisLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 241
Db 303 CAGGAGAGAGAGAGATCTGCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 362
Qy 242 GluSerLysCysArgGluLysGlnArgHisAsnArgSerLysLysGluGluValGln 261
Db 363 GAGAGTCTGTGCGGAGGAGTGCACGCGACACACCGGTCCCTGAGAGAGAGAGAGAGAG 422
Qy 262 ArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeu 281
Db 423 CGAGCCCGTGCAG 482
Qy 282 AsnAspLysGlnLysMetGluGlnHisAsnGluArgAsnSerLysLysLysLysLysLys 301
Db 483 AATGATTCAGTCTGAGTGAACAGGATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 542
Qy 302 AsnMetGluLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 321
Db 543 AATATGAGCTAGCCGAGAGGCTCAAGAGAGTGTATCGAGCAATACAGAGCTTCGTGAGG 602
Qy 322 HisLysAspLysValPheLysHisLysAspLysGlnGlnLysLysLysLysLysLysLys 341
Db 603 CATATCGACAAAGTCTTCAACATAGGAGCTGCGAGAGAGAGAGAGAGAGAGAGAGAG 662
Qy 342 GlnGlnAlaGlnGluMetLysLysLysLysLysLysLysLysLysLysLysLysLysLys 361
Db 663 CAGCAGGCCCGAGAGATGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 722
Qy 362 LeuLysLysGluAlaValGluSerGlnArgMetCysGluLysLysLysLysLysLysLys 381

Db 723 CTCCTGAAGAGAGCGGTGGAGTCCAN-AGGATGTGGAGTGTATGAGCAGAGAGACC 781
Qy 382 HisLeuLysGlnGlnLeuAlaLeu 389
Db 782 CACTA-AAGCAGCAGCTCGCCCTG 804
RESULT 9
BG827157
LOCUS 602751143F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903690 5',
DEFINITION mRNA sequence.
ACCESSION BG827157
VERSION 1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 932)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: King Hong/Rubin Laboratory
CDNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM802 row: f column: 11
High quality sequence start: 6
High quality sequence stop: 775.
Location/Qualifiers
1. 932
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4903690"
/tissue_type="rhabdomyosarcoma"
/lab_host="DRI0B (phage-resistant)"
/clone_lib="NIH MGC 17"
/notes="Organ: muscle; Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAGAG(G). Size-selected for average insert size 1.8kb. Library constructed by King Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

FEATURES
source

Alignment Scores:
Pred. No.: 1-7e-70 Length: 932
Score: 1250.50 Matches: 274
Percent Similarity: 95.82% Conservative: 1
Best Local Similarity: 95.47% Mismatches: 6
Query Match: 44.90% Indels: 9
DB: 12 Gaps: 0
US-10-023-523-44 (1-546) x BG827157 (1-932)

ORIGIN

Qy 257 GluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHis 276
Db 6 GAAGAGGTGTGAGCGGGCCCGGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 65
Qy 277 PheGlnValThrLeuAsnAspLysGlnGlnMetGluGlnHisAsnGluArgAsnSer 296
Db 66 TTCAGGTGACACTGATGACATTCAGCTGCAGATGGACACAGCAATGAGCGCAATCC 125
Qy 297 LysLeuArgGlnGlnAsnMetGluLeuAlaGluArgLeuLysLysLysLysLysLys 316

Db 373 CTGCGTCAAGACAAGTTGAGCTTCTGATCGACTTAAGAAGCTTATCGACCAATATGAG 314
 Qy 318 LeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnGluVal 337
 Db 313 CTAAGAGAAGAGCACATTGACAAAGTTTCAACATAAGGATCTTCAACAGCAACTGGTG 254
 Qy 338 AspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluGluArgHisGlnArg 357
 Db 253 GATGCGAAGCTTCAGCAAGCAACAAGATCTGGAAGAAGTGGAGAGCGTCAACAGCGG 194
 Qy 358 GluLysAspPheLeuLysGluAlaValGlnSerGlnArgMetCysGluLeuMetLys 377
 Db 193 GAAAGGAGTTCTGTTGAAGAGCAGCGATCGAATCCAGCGCATGTGTGAACCTAATGAAA 134
 Qy 378 GlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyThrGluLysPheGluGluPhe 397
 Db 133 CAGCAAGAGACCCACCTCAACAACAGTTAGCACTGTATACAGAGAAAGTTTGAGGAGTTT 74
 Qy 398 GlnAsnThrLeuSerLysSerGluValPheThrThrPheLysGlnGluMetGluLys 417
 Db 73 CAACAACCTTGTCTAAACAAGTGTAGGTTTTCACCACTTTTAA-CAAGAATGCAAAA 15
 Qy 418 MetThrLysLys 421
 Db 14 AAAAAAAAAAAAA 3

RESULT 11
 CB247992 779 bp mRNA linear EST 09-JUL-2003
 LOCUS UI-M-F10-ceb-c-22-0-UI.r1 NIH BMAP_F10 Mus musculus cDNA clone
 DEFINITION IMAGE:6837167 5', mRNA sequence.
 ACCESSION CB247992
 VERSION CB247992.1 GI:28369636
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 779)
 REFERENCE NIH (MGC http://mgi.nci.nih.gov/).
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgaabs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library prepared by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.
 Location/Qualifiers
 1..779
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6837167"
 /tissue_type="whole brain"
 /dev_stage="embryo 12.5dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH BMAP_F10"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with NotI and then cloned

directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is CAGCCAGCAGC. This library was created for the University
 of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."

ORIGIN

Alignment Scores: 4,44e-70 Length: 779
 Pred. No.: 1242.50 Matches: 252
 Score: 97.69% Conservative: 2
 Percent Similarity: 96.92% Mismatches: 5
 Best Local Similarity: 44.61% Indels: 2
 Query Match: 14 Gaps: 1
 DB: 1
 US-10-023-523-44 (1-546) x CB247992 (1-779)
 Qy 119 ArgAsnGlyGluProGluPro---ThrProValValAsnGlyGluLysGluProSerLys 137
 Db 1 AGCAATGGGGAGGCTGAACCCAGGCATTCCAGTCTCAACGGCGAGAAGGAGACCTCTAAG 60
 Qy 138 GlyAspProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAspHisAsq 157
 Db 61 GGAGAGCCTGGAACAGAGGAGATCCGAGCGAGTGATGAGTTGGAGACCGAGACCATCGG 120
 Qy 158 ArgProGlnGluLysLysLysAlaLysGlyLeuGlyLysGluLeuThrLeuLeuMetGln 177
 Db 121 AGCCACAGAGAGAGAGAGCAAGGCTTAGGAGAGGAGATCACTCTGCTGATGCAG 180
 Qy 178 ThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysTyAla 197
 Db 181 ACACCTGAACACGCTGAGTACCCAGAGGAGAGCTGGCTGCACCTGTGCAAGAGTATGCT 240
 Qy 198 GluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGln 217
 Db 241 GAGCTGCTGGAAGAGCATCGAACTCGCAGAGCAGATGAAGCTCTGCAGAGAGAGCAG 300
 Qy 218 SerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAla 237
 Db 301 AGCCAGCTCTGCAGAGAGAGGAGACCATCTCCGAGGGGAAACACAGCAAGGCTGTCTGGCC 360
 Qy 238 ArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGlu 257
 Db 361 CGAAGCAAGCTTGAGAGTCTGTCCCGGAGCTCCACAGCGCACACCGGTCCTGAGAGAA 420
 Qy 258 GluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPhe 277
 Db 421 GAAGCGCTGCAGCGAGCCCGTGAGGAGGAGAGAGCGCAAGAAAGTGACTTCCACTTC 480
 Qy 278 GlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLys 297
 Db 481 CAGGTGACCTGAATGACATTCAGCTGCAGATGGAACAGCATTAACAGCGCAAACTCCAAG 540
 Qy 298 LeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuLeuGluGlnTyArg 317
 Db 541 CTGCGCCAGAGAGATATGGAGCTAGCCGAGAGGCTCAAGAAGTTGATCGAGCAATACGAG 600
 Qy 318 LeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeuVal 337
 Db 601 CTTCTGAGGAGCATATCGCAAAAGTCTTCAACATAAGGACCTGAGCAGCAGCTAGTG 660
 Qy 338 AspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGlnArg 357
 Db 661 GACGCCAAGCTCCAGCAGCGCCAGGAGATCTGAAGAGGAGCAGAGGCGCCACCGAGA 720
 Qy 358 GluLysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLys 377
 Db 721 GAGAAGGAGTTTCTCTCTGAAGGAGCGGTGGAGTC-CAGAGGATGTCGAGGAGCTGATGAAG 779

RESULT 12
 BQ714219

LOCUS BQ714219 942 bp mRNA linear EST 16-JUL-2002
 DEFINITION AGENCOURT 8291711 NIH_MGC_129 Mus musculus cDNA clone IMAGE:6308909
 5' mRNA sequence.
 ACCESSION BQ714219
 VERSION BQ714219.1 GI:21853118
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 942)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgabbs@mail.nih.gov
 Tissue Procurement: Susan L. Sullivan, PhD.
 cDNA Library Preparation: ResGen, Invitrogen Corp
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LRAM13728 row: m column: 06
 High quality sequence stop: 724.
 Location/Qualifiers
 1..942
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:6308909"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 129"
 /note="Organ: olfactory epithelium; Vector: pCMV-SPORT6.1;
 Site 1: EcoRV; Site 2: NotI; Cloned unidirectionally.
 Primer: Oligo dT. Average insert size 2.2 Kb. Constructed
 by ResGen, Invitrogen Corp. Note: this is a NIH_MGC
 Library."

FEATURES

source

ORIGIN

Alignment Scores:
 Pred. No.: 1,966-69 Length: 942
 Score: 1234.00 Matches: 256
 Percent Similarity: 87.58% Conservative: 5
 Best Local Similarity: 85.91% Mismatches: 34
 Query Match: 44.31% Indels: 4
 DB: 13 Gaps: 2

US-10-023-523-44 (1-546) x BQ714219 (1-942)

QY 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerPro 20
 DB 55 ATGAAGAACCAAGACAAAGAAATGGGCTGCTGCAACACTCCAACTCGAAGGCGAGCCG 114
 QY 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
 DB 115 GGGCAACGGAGACAGGACCGAGGAGCCCATGATGACCAAGCCCAACAGACAGCTCCCTGG 174
 QY 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
 DB 175 GCGGAAGCCGAAGGT---TCCACAGCCAGGCTCTCTGGAGAGACCGAGGGGCTCGAGCT 231
 QY 61 ArgThrAlaGlnSerGlyAlaAlaGluArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
 DB 232 AAAGCAGCTCAGCCCTGGGGCCCTCTGTGACGCTCTCTGAGGAGCTGAGCCGCGACGTTGAA 291
 QY 81 AspileLeuSerThrTyrcysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAla 100
 DB 292 GACATCCTGAGTACATAGTGTGGACACAAATCAGGAGGCCCGCTGAGGAGGAGCA 351
 QY 101 GlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrcysValAlaAlaArgAsn 120
 DB 352 CAGGCTGAGCCCACTGAGCCCGAAGACACGAGAGAGTCCCGAAGCTATGACGCCAGGAAT 411

QY 121 GlyGluProGluPro---ThrProValValAsnGlyGluLysGluProSerLysGlyAsp 139
 DB 412 GGGGAGCCTGAACAGGACATTCCTCTCAACGGCGAGAGGAGACCTCTTAAGGAGAG 471
 QY 140 ProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAspHisArgArgPro 159
 DB 472 CCTGGAACAGAGGAGATCCGAGCGAGTGTATGAAGTTGAGACCGAGACCATCGGAGGCCA 531
 QY 160 GlnGluLysLysAlaLysGlyLeuGlyGluLeuThrLeuLeuMetGlnThrLeu 179
 DB 532 CAGGAGAGAGAAAGCCAGGGTCTAGGAAGAGATCACTCTGCTGATGACACACTG 591
 QY 180 AsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysLysLysLysLysLys 199
 DB 592 AACACGCTGAGTACCCAGAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 651
 QY 200 LeuGluGluHisArgAsnSerGlnLysGlnMetLysLysLysLysLysLysLysLys 219
 DB 652 CTGGAAGAGCATCGAACTCGAAGAGCATGAAGCTCTCTGCAAGAGAGAGAGAGAGAG 711
 QY 220 LeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysLysLysLysLysLysLys 239
 DB 712 CTGTCGAGGAGAGAGAGACCATCTCGAGGGGAAACACAGCAGGCTGCTGCTGCTGCTG 771
 QY 240 LysLeuGluSerLeuCysArgGluLeuGlnArgHisLysLysLysLysLysLysLys 259
 DB 772 AAGCTTGAGAGTCTGTGCGGAGCTGCAACGGGCAACACCGGCTGCTGCTGCTGCTGCTG 831
 QY 259 yValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnVa 279
 DB 832 CTGTCGAGGAGCCCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 890
 QY 279 lthLeuAsnAspLysGlnLeuGlnMetGluGlnHisAsnGluArgAsnSer 296
 DB 891 GACACTGAATGACATTCAGCTGCGAGTGAACAGCAGCAGCAGCAGCAGCAGCAGCAGC 942

RESULT 13
 AL961454
 LOCUS AL961454
 DEFINITION AL961454 XGC-gastrula Silurana tropicalis cDNA clone Tgas127p15 5',
 mRNA sequence.
 ACCESSION AL961454
 VERSION AL961454.2 GI:38703803
 KEYWORDS EST.
 SOURCE Silurana tropicalis (western clawed frog)
 ORGANISM Silurana tropicalis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 Xenopodinae; Silurana.
 1 (bases 1 to 945)
 Croning, M.D.R., Ashurst, J.L., Taylor, R., Zorn, A.M. and Rogers, J.
 Sanger Xenopus tropicalis EST project 2001 (11_2003)
 Unpublished (2003)
 On Nov 27, 2002 this sequence version replaced gi:25785049.
 CONTACT: Taylor R
 Sanger Institute
 Hinxton, Cambridgeshire, CB10 1SA, UK
 Email: trop@sanger.ac.uk
 Sanger Xenopus tropicalis EST project 2001
 TROPICALIS_SEQUENCE_ID: Tgas127p15.plkxsp6
 Sequencing primer: SP6
 This sequence is from a Xenopus Gene Collection (XGC) library
 constructed by Aaron M. Zorn.
 cDNA was oligo dT primed from sug of poly A+ RNA from stages 10-13
 gastrulae. EcoRI-NotI cut cDNA was then ligated into pCS107 with
 EcoRI at the 5' end and NotI at the 3' end.
 Vector: pCS107; Site 1: EcoRI; Site 2: NotI
 Host: Escherichia coli XL1-blue.
 Location/Qualifiers
 1..945
 /organism="Silurana tropicalis"
 /mol_type="mRNA"

FEATURES

source

/db_xref="taxon:8364"
 /clone="TGasi127p15"
 /dev_stage="gastrula (stages 10.5-12 mixed)"
 /lab_host="Escherichia coli XL1-blue"
 /clone_lib="XSC-gastrula"
 /note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA was oligo dT primed from Sug of poly A+ RNA from stages 10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end."

ORIGIN

Alignment Scores:

Pred. No.: 2,286-69 Length: 945
 Score: 1233.00 Matches: 254
 Percent Similarity: 85.06% Conservativeness: 25
 Best Local Similarity: 77.44% Mismatches: 35
 Query Match: 44.27% Indels: 15
 DB: 9 Gaps: 3

US-10-023-523-44 (1-546) x AL961454 (1-945)

QY 79 LeuGluAspIleLeuSerThrTyrcysValAspAsnAsnGlnGlyProGlyGluAsp 98
 Db 1 CTGGAAGATTTCCTAGTACATATTGTTGCTTACTAGCAAGAGAGGAGGAGACTCG 59
 QY 99 GlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyValAla 118
 Db 60 -----GAGGAGCAGATAGGAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 92
 QY 119 ArgAsnGlyGluProGlu---ProThrProValValAsnGlyGlyGluProSerLys 137
 Db 93 AAAAAATGCTGACTACTAGTACTAGTACTAGTACTAGTACTAGTACTAGTACTAGTACTAGT 152
 QY 138 GlyAspProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAspHisArg 157
 Db 153 GTTTAGAGCAGCTAGGAGCAGCTAGGAGCT-----GAAAGAGCTCAGAA 200
 QY 158 ArgProGlnGlnGlyLysLysAlaGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 177
 Db 201 CGAATGCAAGACAAACAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 260
 QY 178 ThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaAlaLeuCysLysLysTyValAla 197
 Db 261 ACTCTGAACACACTCAGACCCAG 320
 QY 198 GluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuGlnLysLysGln 217
 Db 321 GAACTGTTGGAGGAGCAGGAGCAGGAGCAGGAGCAGGAGCAGGAGCAGGAGCAGGAGCAGGAG 380
 QY 218 SerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAla 237
 Db 381 ACCGAGCTTATCAAGAAAGAGATCTACTTCGTAATGAGCAGCAGGAGGAGGAGGAGGAGGAG 440
 QY 238 ArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGlu 257
 Db 441 AGAAGCAAGCTCGAAGATTGTGACAGAGAGCTCAGAGCAATGAGGAGTACTGCAAAAAAAGAG 500
 QY 258 GluGlyValGlnArgAlaArgGluGluGluLysArgLysGlyValThrSerHisPhe 277
 Db 501 GAAGAGTACAGAGGCTCTGAG 560
 QY 278 GlnValThrLeuAsnAspIleGlnLeuGlnMetGlnHisAsnGlnGluArgAsnSerLys 297
 Db 561 CAAGTACGCTCAATGATCATCCAGTACAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 620
 QY 298 LeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuLeuGlnTyValGlu 317
 Db 621 CTGGTCAAGCAAGATTGAGCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 680
 QY 318 LeuArgGluGluHisLysAspLysValPheLysHisLysAspLeuGlnGlnLeuVal 337
 Db 681 CTAAGAGAGAGGAGCAGTATGCAAAAGTTTTCAAAACATAGGAGATCTTCACAGGAGCACTGGTG 740

QY 338 AspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluLysGlnArg 357
 Db 741 GATGCAAGAGCTTCAGCAAGCAGCAGCAAGAAATGTTGAAGAAGAGTGAAGAGGCTCACCAGCGG 800
 QY 358 GlutylAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLys 377
 Db 801 GAAAGAGAGTTCTGTTGAAGAGGAGGAGTGAATCCAGCGCATGTGTGAATTAATGAAA 860
 QY 378 GlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyThrGluLysPheGluGluPhe 397
 Db 861 CAGCAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 920
 QY 398 GlnAsnThrLeuSerLysSerSer 405
 Db 921 CAAACACACCTTGTCTAAAGCAAT 944

RESULT 14

BU701958

LOCUS

BU701958

DEFINITION

UI-M-F10-bys-i-05-0-UI.r1 NIH_BMAP_F10 Mus musculus cDNA clone

ACCESSION

BU701958

VERSION

BU701958.1

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

1 (bases 1 to 725)

AUTHORS

NIH-MGC <http://mgc.ncbi.nlm.nih.gov/>

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: sga@pshs.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA sequencing by: Dr. M. Bento Soares, University of Iowa
 clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

FEATURES

Seq primer: pyX-5.

Location/Qualifiers

1..725

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:5697868"

/tissue_type="whole brain"

/dev_stage="embryo 12.5dpc"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NIH BMAP F10"

/notes="Organ: Brain; Vector: pyX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Benton and Davis, Molecular Cloning, 2nd ed., Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 1988. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pyX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is 'CAGCCACGAG'. This library was created for the University of Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

ORGANISM:

Db 251 CACAGCAAGGCTGTCTCTGGCCGCAAGCAAGCTTGAGAGTCTGTGCGCGAGCTGCAACGG 310
 Qy 251 HisAenArgSerLeuLysGluGluGluValGlnArgAlaArgGluGluGluLysArg 270
 Db 311 CACAACCCGTCCTTGAGAGAAAGGGGTGCGAGCGCCGTGAGGAGGAGGAGCGC 370
 Qy 271 LysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGln 290
 Db 371 AAAGAAGTGACTTCCAGCTCCAGCTGACACTGAATGACATTCAGCTGCGAGTGGAAACAG 430
 Qy 291 HisAenGluArgAsnSerLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLys 310
 Db 431 CATAACGAGCGAAATCTCAAGCTGCGCAGGAGGAGTATGAGCTAGCCGAGAGGCTCAAG 490
 Qy 311 LysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAspLysValPheLysHisLys 330
 Db 491 AAGTTGATCGAGCAATACGAGCTTCGTGAGGAGCATATCGACAAGTCTTCAAAACATAAG 550
 Qy 331 AspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGlu 350
 Db 551 GACCTGCGAGCAGCTAGTGTGAGCGCAAGCTCCAGCAGCCCGCAGGAGATGCTGAAGGAG 610
 Qy 351 AlaGluGluArgHisGlnArgLysAspPheLeuLeuLysGluAlaValGluSerGln 370
 Db 611 GCAGAGGAGCGGCACCGAGGAGAGAGAGGAGTTCTCTGGAAGCAAGCGGTGAGTCCCAAG 670
 Qy 371 ArgMetCysGluLeuMetLysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyr 390
 Db 671 AGGATGTCGAGCTGATGAGACAGCAGAGACCCACCTAAAGCAGCAGCTGCGCCCTGTAC 730
 Qy 391 Thr 391
 Db 731 ACG 733
 RESULT 16
 CF995807
 LOCUS
 DEFINITION AGENCOURT_16109164 NIH MGC 221 Homo sapiens cDNA clone
 IMAGE:30708497 5', mRNA sequence.
 ACCESSION
 VERSION CF995807
 KEYWORDS
 EST.
 SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: James Martin, University of Iowa
 cDNA Library Prepared by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: NDAMI074 row: m column: 18
 High quality sequence stop: 631.
 Location/Qualifiers
 1..759
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30708497"
 /lab_host="DH10B Tona"
 /clone_lib="NIH MGC 221"
 /note="Organ: mixed; Vector: pYX-Asc; Site_1: ECORI;
 source

ORIGIN

Alignment Scores:
 Pred. No.: 1,09e-67 Length: 759
 Score: 1205.00 Matches: 242
 Percent Similarity: 98.39% Conservative: 2
 Best Local Similarity: 97.58% Mismatches: 4
 Query Match: 43.27% Indels: 0
 DB: 14 Gaps: 0.

US-10-023-523-44 (1-546) x CF995807 (1-759)

Qy 204 ArgAsnSerGlnLysGlnMetLysLeuLysGlnLysGlnSerGlnLeuValGlnGlu 223
 Db 9 CGGAATTCACAGAAAGCAGATGAAGCTCTCTACAGAAAAGCAGAGCAGCTGTGTCAAGAG 68
 Qy 224 LysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSer 243
 Db 69 AAGGACCACCTGCGCGGTGAGCAGACAAAGGCGCTCTGCGCCGCGAGCAAGCTTGAGAGC 128
 Qy 244 LeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAla 263
 Db 129 CTATGCGCTGAGCTGCGCGGCAACACCTCTCTCAAGAAAGAGGTGTGCGCGGCGCC 188
 Qy 264 ArgGluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAsp 283
 Db 189 CGGAGGAGGAGGAGAAAGCGCAAGAGGTGACCTCGCACTTCAGGTGACACTGAATGAC 248
 Qy 284 IleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluGlnMet 303
 Db 249 ATTGAGCTGAGATGAACACAGCAATGAGCGCACTCCAGCTGCGCAAGAGAACATG 308
 Qy 304 GluLeuAlaGluArgLeuLysLysLeuIleGluGlnTyrGluLeuArgGluGluHisIle 323
 Db 309 GAGCTGGCTGAGAGGCTCAAGAAGCTGATTGAGCAGTATGAGCTGCGCGAGGAGCATATC 368
 Qy 324 AspLysValPheLysHisLysAspLeuGlnGlnLeuValAspAlaLysLeuGlnGln 343
 Db 369 GACAAGTCTTCAACACAGAGGACCTTACACAGAGCTGTGTGATGCCAAGCTCCAGCAG 428
 Qy 344 AlaGlnGluMetLeuLysGluAlaGluArgHisGlnArgGluLysAspPheLeuLeu 363
 Db 429 GCCCAGGAGATGTAAAGGAGGAGGAGCGCACCGCGGAGAGGAGTATTTCTCTCTG 488
 Qy 364 LysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLeu 383
 Db 489 AAAGAGCAGTAGAGTCCCGAGAGGATGTGTGAGCTGTGATGAGCAGCAGCAGACCCACCTG 548
 Qy 384 LysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLysSerLys 403
 Db 549 AAGCAACAGCTTCCCTATACACAGAGAAGTTTGGAGGTTCCAGAACACACTTTTCCAAA 608
 Qy 404 SerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLys 423
 Db 609 AGCAGGAGGATTTTCCACATNTCAGCAGGAGATGGAAGAGNATGACTAGAGATCCAG 668
 Qy 424 LysLeuGluLysGluThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeu 443
 Db 669 AAGCTGAGAGAAAGAACCCACCATGTACCGTCCCGTGGGAGGAGCAGCAACAGGCCCTG 728

Site 2: NotI; Library is oligo-dT primed and directionally cloned. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to RNA size fraction, ligated with EcoR I adaptor, digested with Not I and then cloned directionally into pYX-Asc vector. Average insert size 4-5kb. Adaptors 5' (AATCGCACGAGG)3' and 5' (CTCTGCGCG)3'. 3' Linker sequence - GCGGCGGTGAGAGCC T18. Sequencing primers 3' end: T3 promoter primer 5'd (ATTACCTCTCACTAAGGGA)3'. 5' End: T7 promoter primer 5'd (TAATACGACTCACTATAGG)3'. Library was constructed in the laboratory of M. Bento Soares. Note: this is a NIH_MGC Library"

QY 444 LeuGluMetAlaGluGluLysThr 451
 DB |||||
 729 CTTGAGATGGCTGAAGAAAAACA 752

RESULT 17
 AK084639
 LOCUS
 DEFINITION
 Mus musculus 13 days embryo heart cDNA, RIKEN full-length enriched library, clone:D330024K18 product:MUSCLE-DERIVED PROTEIN MDP77 VARIANT 1, full insert sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

AK084639.1 GI:26351202
 HTC; CAP trapper.
 Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Carninci, P., and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 99279253
 10349636

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 20499374
 11042159

3 Shibata, K., Itoh, M., Aizawa, K., Katsuka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Kazama, M., Nishine, T., Harada, A., Yamamoto, K., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Nishikawa, K., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipipillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076661

4 The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6 (bases 1 to 1608)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirakawa, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, D., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
 Direct Submission
 Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC),

RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site for further details.
 URL:http://genome.gsc.riken.go.jp/
 URL:http://fantom.gsc.riken.go.jp/

FEATURES
source

Location/Qualifiers
 1..1608
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM DB:D330024K18"
 /db_xref="MGI:2421329"
 /db_xref="taxon:10090"
 /clone="D330024K18"
 /tissue_type="heart"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="13 days embryo"
 218..>1606
 /notes="unnamed protein product; MUSCLE-DERIVED PROTEIN MDP77 VARIANT 1 (SPTR)AAL33909, evidence: PASTY, 99.1%ID, 100%length, match=2058)
 putative"
 /codon_start=1
 /protein_id="BAC39238.1"
 /db_xref="GI:26351203"
 /translation="MEINHPDOLSVHEPTPPGDSSLSNONGPKQDGRCSGQAPPEQSLHPKGAHDVAELSRQLEDILITVSGSAASPRGKSTSETKEPPNTEAPDNEDVYEETEDIDREPTAPEPAAKPEVSNKEQLEKILKGKZENILMQLNKLQAPEKLDLPFKTTELDEHTEQKLLKQAOQKREKQDQSQSHNRVAVLRSKLESECRETQRNKILKESTLQAREREERKEITSHFTITLDTQTEQSERMRLKQENTLEAKLSIDQVELREHLDIKFKRELQKLYAKLEAEALQMEAEERHREKYLNNQAEWLQAKVLKEQETVLAQUTLYSGRPERFEQSTLTNSNVPATPKQMDKTTKMKLEKDTATKARFENCNKALLDMIEKALRAKEYECFVVKIQRLNCLRALQERKELYKKIRE"

CDS

ORIGIN
 Alignment Scores:
 Pred. No.: 1,86e-66 Length: 1608
 Score: 1192.00 Matches: 248
 Percent Similarity: 68.20% Conservative: 78
 Best Local Similarity: 51.88% Mismatches: 98
 Query Match: 42.80% Indels: 54
 DB: 11 Gaps: 6

US-10-023-523-44 (1-546) x AK084639 (1-1608)
 QY 34 ProSerGlnAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArg 53
 DB |||||
 263 CCCCCAGGGACAGCTCATCTACTCAATCAACCGCCCGGAAAGCAAGATGCGAGCGG 322
 QY 54 LysProGluGlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeu----- 68
 DB |||||
 323 TGCTCAACCTCAGGCCCAAGCG--CCAGAGCAAGAGGGAAGCGCTGCATCCCGAGAGGGA 379
 QY 69 ---ArgAspValSerGluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyCys 87
 DB |||||
 380 GCCCATGATGTCGGGAAGAGTTGACAGCACTGAAGACATCATAGCACGATAT---- 436
 QY 88 ValAspAsnAsnGlnGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluPro 107
 DB |||||
 437 -----GGGTCTGCTGCCAGTCCCCCGAGGGAAG 463
 QY 108 GluAspAlaGluLysSerArgThrTyValAlaArgAsnGlyGluProGluProThrPro 127
 DB |||||
 464 GAGAGCACCTCTGAACACTAAG----- 484

URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222,

Fax: 81-45-503-9216)

COMMENT
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.go.jp/>

URL: <http://fantom.gsc.riken.go.jp/>.

FEATURES

source

```
1. 2480
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:D830027D20"
/db_xref="MGI:2423165"
/db_xref="taxon:10090"
/clone="D830027D20"
/tissue_type="heart"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="16 days neonate"
1. 2480
/misc_feature
/notes="MUSCLE-DEIVED PROTEIN MDP77 VARIANT 1
(SPTR|AAL33909, evidence: FASTV, 99.1%ID, 100%length,
match=2058)"
```

ORIGIN

Alignment Scores:

```
Pred. No.: 3.67e-66 Length: 2480
Score: 1191.00 Matches: 262
Percent Similarity: 63.47% Conservative: 82
Best Local Similarity: 48.34% Mismatches: 134
Query Match: 42.76% Indels: 64
Db: 11 Gaps: 8
```

US-10-023-523-44 (1-546) x AK085904 (1-2480)

```
QY 34 ProSerGlnAlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArg 53
DB 156 CCCCAGGGGAGAGTCTACCTCAATCAAAACGCCCGGAAACAGCAGATGGCGGCG 215
QY 54 LysProGluGlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeu- 68
DB 216 TGCTCAACTCAGGCCAAGCG---CCAGAGCAAGAGGGAAGCTGCTCCCGAGAGGGA 272
QY 69 ---ArgAspValSerGluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyCys 87
DB 273 GCCCATGATGTCGCGGAGAGATTGAGCAGGCAACTGGGAAGACATCATAAAGCAGCTAT 329
QY 88 ValAspAsnAsnGlnGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluPro 107
DB 330 ---GGTCTGCTGCCAGTCCCGAGGGAAG 356
QY 108 GluAspAlaGluLysSerArgThrTyrValAlaArgAsnGlyGluProGluProThrPro 127
DB 357 GAGACACCTCTGAACATAAG--- 377
QY 128 ValValAsnGlyGluLysGluProSerLysGlyAspProAsnThrGluLeuLe- 145
DB 378 ---GAGACCCCCCAACACACAGAGCCACCAACATGAGGATGTTGACTAT 425
QY 146 ArgGlnSerAspGluValGlyAspArgAspHisArgArgProGln- 160
DB 426 GAAGAAACCACTGAAGAGATAGACAGAGAACCACTGCTCTGAAGAGCCAGCCGCGCAGCC 485
QY 161 ---GluLysLysLysAlaLysGlyLeuGly 169
DB 486 AAAGAGCTGTGACGAATAAGAGCAAAAGCTGGAAAGAAATCTTAAAGGATTAGGT 545
QY 170 LysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeu 189
DB 546 AAAGAAGCTAATTGCTCATGTCAAAACCTGAACAGCTGCAAGCAGCTTGAAGAAAGGCTT 605
```

```
QY 190 AlaAlaLeuCysLysLysTyrAlaGluLeuLeuGluGluHisArgAsnSerGlnLys-GI 209
DB 606 GATTTTATTCAAGAGTATCTGTAATTCCTGATGAACATCCACCCAGCAAGAGAAA 665
QY 209 nMetLysLeuLeuGlnLysGlnSerGlnLeuValGlnGlnLysAspHisLeuArgGI 229
DB 666 GCTGAGCTCTCTACTGAACAGCAGCGCGCAGACCCAGAGAGAGAGAGCAGTTGCAGAG 725
QY 229 YGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGI 249
DB 726 TGAGCAACACAGAGCGCTCTGCTCGAACCAACTCGAGAGTCTGTGTCGGAGCTACA 785
QY 249 nArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGluGluGluLys 269
DB 786 GAGCACAACAGAGCGCTGAAGGAAGAACCTTCCAAACGGCGCGCGAGGAAGAGAA 845
QY 269 sArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGI 289
DB 846 GAGGAAGAGATCACAGTCACTTTCAACTACCTGACAGATATCCAGACTCAGATCGA 905
QY 289 uGlnHisAsnGluArgAsnSerLysLeuArgGlnGlnLeuMetGlu-LeuAlaGluArgL 309
DB 906 ACAACAGAGTGAAGCAATATGAAGCTCTGCCAGGAGAACACACAGAGCCTTGCAGAAAAAC 965
QY 309 euLysLysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAspLysValPheLysH 329
DB 966 TGAAGAGCATCATGACCATGACAGTACAGCTCAGAGAGGAGCATCTGGACAAATATTCAAAC 1025
QY 329 isLysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGlnMetLeuL 349
DB 1026 ATCGAGAGTGCAGCAGAGAGCTGTGATGCGAACTTGAAGAGGCTCAAGAACTGATGC 1085
QY 349 ysGluAlaGluGluArgHisGlnArgGluLysAspPheLeuLeuLysGluAlaValGluS 369
DB 1086 AGGAGCAGAGAGCGGCACAGACGCGAGAGGAATATTGCTGTAACACCGCGCGAGAGT 1145
QY 369 erGlnArgMetCysGluLeuMetLysGlnGlnGlnThrHisLeuLysGlnGlnLeuAlaL 389
DB 1146 GGAACCTTCAGGCCAAGAGTCTGAAGAGGAGCAGAGAGCGGTGCTGCAGGCTCAGCTC 1205
QY 389 euTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluValPheT 409
DB 1206 TCTACTCAGGAAGGTTTGAAGAGTTCAGAGTACACTGACAAAGCAATGAAGTCTTTG 1265
QY 409 hrThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLysLysLysLysL 429
DB 1266 CCATTTCAAGCAGGAATGGAATGACACAACTAAGAAATGAAGAAGCTGGAGAGGACA 1325
QY 429 hrThrMetTyrArgSerArgTrpGluSerSerAsnLysAlaLeuLeuGluMetAlaGluG 449
DB 1326 CCGCCAGCTGGAAGCCGATTTGAGAAATTTGAATAATTGTAACAAAGCCCTTTGGACAT 1385
QY 449 lLysThrValArgAspLysGluLeuGluGlyLeuGlnValLysIleGlnArgLeuGluL 469
DB 1386 AGAAACCTTCGAGAGTAAAGATTCGAGTGTTCGTGATGAATAATCCAAAGCTGGAGA 1445
QY 469 ysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgValGlnAspLys 489
DB 1446 ACCTGTGTCGTCTTACAGAGCAGAGAGAAAGAACTTACAGAGAAATCAGAGAGCAA 1505
QY 489 erAlaGlyGlnGlnGlySerLeuThrAspSerGlyProGluArgArgProGluGlyProG 509
DB 1506 AAATGTCCGAAAAAGAGACCAAGTTTCAGCGCACCTCTGAGGAAGAGCCAGAG---CCAA 1562
QY 509 lyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyrProGlyAlaProS 529
DB 1563 GCGTCTCTGAATGAAGAGGTGGAGCCAGAGAGCCAAACAGTCTTTCAAAAAGCTGTGG 1622
QY 529 erThrGluAlaSerGlyGlnThr- 1622
DB 1623 AAAACCTGGCCACAGCTTTTACCATCTCTCATCCAGAGTTTCAACCCAGACCAACCAA 1682
```

```

QY      542 hr 542
Db      1683 CG 1684

RESULT 19
BU849604
LOCUS   BU849604
DEFINITION AGENCOURT_10437001 NIH_MGC_109 Homo sapiens cDNA clone
IMAGE: 6598234 5', mRNA sequence.
ACCESSION BU849604
VERSION   BU849604.1 GI:24034567
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 799)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM2830 row: 1 column: 10
High quality sequence stop: 565.
Location/Qualifiers
1..799
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6598234"
/tissue_type="teratocarcinoma, cell line"
/lab_host="PH10B (phage-resistant)"
/clone_lib="NIH MGC 109"
/note="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using Zap-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

ORIGIN
Alignment Scores:
Pred. No.: 2,04e-65 Length: 799
Score: 1170.00 Matches: 239
Percent Similarity: 95.79% Conservative: 2
Best Local Similarity: 95.98% Mismatches: 4
Query Match: 42.01% Indels: 5
DB: 13 Gaps: 2

US-10-023-523-44 (1-546) x BU849604 (1-799)

QY      1 MetLysAsnGlnAspLysGlnGlnSerAsnProLysSerPro 20
Db      41 ATGAAGACCAAGACAAAGACCGGGCTGCAATCAATCCAAAGACGCCA 100
QY      21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
Db      101 GGACACCGGACGAGCAGGCCCGGAGGAGCCAGGAGCGGCCGAGCGGCTCTGCA 160
QY      41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
Db      161 GTAGAAAGCAGAGAGTCCCGGACGACCCAGGCTCTCCGGAAGCCGCGGGGCTCAAGCC 220
QY      61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80

```

```

Db      221 AGAACGGCTCAGTCTGGGGCCCTTCGTGATGCTCTGAGAGCTGAGCGCAACTGGAA 280
QY      81 AspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAla 100
Db      281 GACATACTGAGCACATACCTGCTGGGAGATAACCGAGGGGGCCCGGCGAGATGGGCA 340
QY      101 GlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsn 120
Db      341 CAGGGTGAGCGGGCTGAACCGAAGATGCGAGAGTCCCGGACCTATGTCGCAAGGAT 400
QY      121 GlyGluProGluProThrProValValAsnGlyGluLysGluProSerLysGlyAspPro 140
Db      401 GGGGAGCCTGAACCAACTCCAGTAGTCAATGAGAGAGAGAAACCTCCAAAGGGGATCCA 460
QY      141 AsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHisArgArgProGln 160
Db      461 AACACAGAGAGATCCGGCAGAGTGCAGAGGTGCGAGAGCCGAGACCATCCAAAGGCCACAG 520
QY      161 GluLysLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsn 180
Db      521 GAGAGAGAAAAGCCCAAGGTTTGGGAGGAGAGACGCTTGTGATGTCAGACATTAAT 580
QY      181 ThrLeuSerThrProGluGluLysLeuAlaAlaLeuCysLysLysLysTyrAlaGluLeu 200
Db      581 ACTCTGAGTACCCAGAGAGAGAGCTGGGTGCTCTGTGCAAGAGATGCTGAACCTGCTG 640
QY      201 GluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeu 220
Db      641 GAGAGACCCCGAATTCACAGAGCAGATGAGTCTCTACAGAAAAGCAGAGCCAGCTG 700
QY      221 ValGlnGluLysAspHisLeu---ArgGlyGluHisSerLysAla-ValLeu-AlaArgS 239
Db      701 GTGCAAGAGAAAGACCACTGNTGCGGNT-GAGCACAGACAGCCGGTCTCTGGGCGCGCA 759
QY      239 erLysLeu---GluSerLeuCys 245
Db      760 GCAAGCTTGGAGAGCCCTATGTC 782

RESULT 20
CA327097 750 bp mRNA linear EST 09-JUL-2003
LOCUS   UI-M-FYO-cv-g-01-0-UI.r1 NIH BMAP_FYO Mus musculus cDNA clone
DEFINITION IMAGE:6826106 5', mRNA sequence.
CA327097
VERSION   CA327097.1 GI:24545195
KEYWORDS  EST.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 750)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
this clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
Location/Qualifiers
1..750
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"

FEATURES
source

```

[illegible]

Db 421 GAAGTGTTCACCACTTCAACAGGAGATGGAAGACATGACAGAGATCAAGAGCTG 480

Qy 426 GtlyysGluThrThrMetTyrArgSerArgTrrGluSerSerhshnlysalaleuLeuGlu 445

Db 481 GAGAGAGAAACCAACCATGATTCATCCCGTGGGAAAGCAGCAACAAAGGCTCTGCTGGAG 540

Qy 446 MetAlaGluGluLysThrValArgAspLysGluLeuGluGluLysVallystleGln 465

Db 541 ATGGCTGAAGAGAAACCGTCCGGACAAAGAGCTGGAGGGCTGAGGTGAATCCAG 600

Qy 466 ArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnlyArgVal 485

Db 601 CGGCTGGAGAAGCTGTCCGAGCACTGCAGACCGCAATGACCTCAACAAGAGGTA 660

Qy 486 GlnAspLeuSerAlaGlyGlnGlySerLeuThrAspSerClyProGluArgPro 505

Db 661 CAGCACTGACTCAGAGGGG-----ATCACTGACATTGGCTCTGAGCGGAGGCA 711

Qy 506 Glu 506

Db 712 GAG 714

RESULT 23

BF783468
LOCUS 602111238F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:4239513
DEFINITION 5', mRNA sequence.

ACCESSION BF783468

VERSION BF783468.1 GI:12088504

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 819)
NIH-MGC <http://imgc.ncl.nih.gov/>

AUTHORS

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM9852 row: 1 column: 10
High quality sequence stop: 758.

FEATURES

source

1..819
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4239513"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Kid14"
/note="Organ: Kidney; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."|

ORIGIN

Alignment Scores:

Pred. No.: 5,036-64 Length: 819
Score: 1148.50 Matches: 240
Percent Similarity: 92.05% Conservative: 3
Best Local Similarity: 90.91% Mismatches: 17
Query Match: 41.24% Indels: 5
DB: 10 Gaps: 1

US-10-023-523-44 (1-546) x BF783468 (1-819)

Qy 75 LeuSerArgGlnLeuGluAspIleuSerThrTyrCysValAspAsnGlnGly 94

Db 2 CTGAGCCGCGAGTTGGAAAGACATCTGTAGTACATCTGTGTGGCAACAATCAGGAGGC 61

Qy 95 ProGlyGluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArg 114

Db 62 CCGGCTGAGGAGGAGGACACAGGGTGAGCCCACTGAGCGGGAAGACACGGAAGTCCCGA 121

Qy 115 ThrTyrValAlaArgAsnGlyGluProGluPro---ThrProValValAsnGlyGluLys 133

Db 122 ACCTATGCAAGCAGCAATGGGAGCCCTGAACAGGCACTTCAGTCTCAACGGCGAGAG 181

Qy 134 GluProSerLysGlyAspProAsnThrGluGluLeuArgGlnSerAspGluValGlyAsp 153

Db 182 GAGACCTCTAAGGAGAGACCTGGAACACAGAGAGATCCGAGCAAGTATGATGAGGTTGGAGAC 241

Qy 154 ArgAspHisArgArgProGlnGlnLysLysLysAlaLysGlyLeuGlyLysGluIleThr 173

Db 242 CGAGACCATCGAGGCCACAGGAGAGAGAGAAAGCAAGGGTCTAGGAGAGAGATCACT 301

Qy 174 LeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCys 193

Db 302 CTGCTGATGACAGACACTGAACACGCTGAGTACCCAGAGGAGAGCTGGCTGCATCTGCG 361

Qy 194 LysLysTyrAlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeu 213

Db 362 AAGAAGTATGCTGAGCTGCTGGAAGAGATCGGAACCTCGAGAAAGCAGATGAAGCTCTG 421

Qy 214 GlnLysLysGlnSerGlnLeuValGlnGlnLysAspHisLeuArgGlyGluHisSerLys 233

Db 422 CAGAAGAAGCAGAGCCAGCTCGTCAGAGAGAGAGACCATCTCGAGGGGGAACACAGCAAG 481

Qy 234 AlaValLeuAlaArgSerLysLeuLeuSerLeuCysArgGluLeuGlnArgHisAsnArg 253

Db 482 GCTCTCTGCGCCGAGCAAGCTTGAGAGTCTGTGCGCGGAGCTGCAACGCGCACACCGG 541

Qy 254 SerLeuLysGluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluVal 273

Db 542 TCCCTGAAGGAGAGAGCGCTGCAGAGAGCCGAGAGAGGAGGAGAGAGAGAGAGAGAGTG 601

Qy 274 ThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGlu 293

Db 602 ACTTCACATCTCCAGGTGACACTGAATGACATTCAGCTGCAGATGGAACAGCATAACGAG 661

Qy 294 ArgAsn-SerLysLeuArgGlnGluAsnMetGluLeuAlaGlu-ArgLeu-LysLysLeu 312

Db 662 CGAAACCTCCAGCTGGCCGAGGAGATATGGAGCTAGCCGAGAGGCTCAAGAAGATTG 721

Qy 313 IleGluGlnTyrGluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeu 332

Db 722 ATGAGCAATACGAGGCTTCGAGGAGCTTTTCGACAAAGTCTT-AAACATTAGGACTGC 780

Qy 333 GlnGln 334

Db 781 GAGCAG 786

RESULT 24

AK031023

LOCUS

DEFINITION

AK031023 4452 bp mRNA linear HTC 18-SEP-2003
Mus musculus adult male thymus cDNA, RIKEN full-length enriched
library, clone:5830498L23 product:hypothetical protein, full insert
sequence.

ACCESSION AK031023

VERSION AK031023.1 GI:26082107

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

1 Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

```

JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagao, K., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Katsunuma, T., Tashiro, H., Itoh, M.,
Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Tozawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--394-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
MEDLINE 11076861
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
MEDLINE 11076861
REFERENCE 6
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration and Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
URL: http://location.qualifiers
FEATURES source
1. 4452
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="PANTOM DB:5830498L23"
/db_xref="MG:2394192"
/db_xref="taxon:10090"
/clone="5830498L23"

```

```

/sex="male"
/tissue_type="thymus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
1. 4452
/misc_feature
/notes="hypothetical protein (evidence:
rscds, ProCrest, decoder, NCBI CDS Predictor, Longest-ORF)"
ORIGIN
Alignment Scores:
Pred. No.: 2,35e-62 Length: 4452
Score: 1136.50 Matches: 260
Percent Similarity: 43.97% Conservative: 68
Best Local Similarity: 34.85% Mismatches: 67
Query Match: 40.81% Indels: 351
DB: 11 Gaps: 5
US-10-023-523-44 (1-546) x AK031023 (1-4452)
QY 35 SerGlnAlaProAlaValGluAlaGluGlyPro-----GlySerSerGlnAlaPro 52
DB 958 TCTCAAGCAATGATATCTTCAACATCAAGACCCAGTTGGTGGCACAACAAAGAAA 1017
QY 53 ArgLysProGluGlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSer 72
DB 1018 CATTCACCTGGAAGGGGATGAAGC-----AGT 1044
QY 73 GluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysVal---AspAsnAsn 91
DB 1045 GACTTATTATCAAGAAACAGAAATTTGGTGAAGCTAGTATCTGTACACAGAGAAAGAA 1104
QY 92 GlnGlyGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGlu 111
DB 1105 GAAGAAATTCCTGGAGGAGGAGCTCGAAGAGTCTCTCTGATGGCCAGCAAGATTACAG 1164
QY 112 LysSerArgThrTyrValAlaArgAsnGlyGluProGluProThrProValValAsnGly 131
DB 1165 TGACAGCAGG----- 1173
QY 132 GluLysGluProSerLysGlyAspProAsnThrGluGluIleArgGlnSerAspGluVal 151
DB 1173 ----- 1173
QY 152 GlyAspArgAspHisArgArgProGlnGluLysLysLysLysLysLysLysLysLys 171
DB 1174 -----AAACAAGAGAGAGACCTTAGGAAAGAA 1200
QY 172 IleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaAla 191
DB 1201 GTTTTATTACTGATGCAAGCGCTAAACACCCCTTCAACCCAGAGAGAGAGTGCACCT 1260
QY 192 LeuCysLysLysTyrAlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLys 211
DB 1261 CTCTGTGAAGAAATATGCTGATCTCTCTGGAGAAACACGAGAAATGTTTCAGAAAACAA 1320
QY 212 LeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHis 231
DB 1321 ATTCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
QY 232 SerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGluArgHis 251
DB 1381 AGCAAGGCATCTTTGGCAAGAGCAAACTGGGAATCTCTTTGAGGAGAACTTCAGCGTCA 1440
QY 252 AsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGluGluGluGluArgLys 271
DB 1441 AATAAGACCTTAAAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
QY 272 GluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuMetGlnGlnHis 291
DB 1501 GAGGCAACAGACATTTCCAGATAACTCTAAATGAAATCCCAAGCTCAGTTGGACAAACAT 1560
QY 292 AsnGluArgAsnSerLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLys 311
DB -----

```


Db 1561 GACATCCACATCCCAACTCGCAGGAGAACATTTGAACCTGGGAGAGAGTTGAAGAG 1620
 QY 312 LeuileGluGlnTyrGluLeuArgGluGluHisIleAspLysValPheLysHisLysAsp 331
 Db 1621 CTTATTGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
 QY 332 LeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAla 351
 Db 1681 TTGCACACACAGCTGTGTGATGCCAACTTCAGCAACACACAGCTGATATAAAGAGCT 1740
 QY 352 GluGluArgHisGlnArgLysAspPhe 361
 Db 1741 GATGAAAAACATCAGAGAGAGAGAGTTGTGAAGTTCTGTTCTTTAAATAAAGGAA 1800
 QY 361 361
 Db 1801 AGGTATTGCCATATTAACTAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
 QY 361 361
 Db 1861 GGAAGCTTCTGGCTACTTATTCTATTTCAGAGATCCACATGCTGCTGCTGCTGCTGCTGCTGCT 1920
 QY 361 361
 Db 1921 TTAATAAAGCATGCTATACCATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
 QY 361 361
 Db 1981 GGAATGATTTTCAATGCTGTAATTTAAATTTTATGATGATGATGATGATGATGATGATGATGAT 2040
 QY 361 361
 Db 2041 TTGTTGAGTCAAGGCTCTCATATTGATGCTCTGGCTGGGATGACCTTGAACACTACTTCAA 2100
 QY 361 361
 Db 2101 GCACATCCAAATTTAGAGAGCTCTGGGATTAACCTGGCTTCTACTTACATATA 2160
 QY 361 361
 Db 2161 AAGCAAGCACATCTCCAGTCTCACACCCATTTGTTGTTCTCTAGACTTACTACTTT 2220
 QY 361 361
 Db 2221 CTTCCATAGCTTTCCGCCCATCTTCTCTGTTCTCTAGTAGTAGAGAGTTTCTCAG 2280
 QY 361 361
 Db 2281 TCATGTAGAGGTACTTAGCCATGTTTCTGCTAATATACACATGATGATTTTTCACC 2340
 QY 361 361
 Db 2341 TGAACACCAACCAACACCATCTTTAAACCTCAACTTTGTAAATGTTTAGGCTCCAC 2400
 QY 361 361
 Db 2401 TATACCTTTTACTTAAGCAGTAAATAATATAAGTATTCAAACATTTCTACTTCTAA 2460
 QY 361 361
 Db 2461 TGATGGAAGTATTGTAAATGTTAACTGTTGTAATAAGGTTTCCATTTTGGAGGTA 2520
 QY 361 361
 Db 2521 GTCATTTCTGCTAAATAAGTCAAGCTGAAGGATGTGGAGTAACATGCTGCTGCTGCTGCTGCTGCTGCT 2580
 QY 362 362
 Db 2581 TATGTTCTTATTGTAATGCTATTAGTACTTTTGCAGGATGTTTCTATTTTATTGTTA 2640
 QY 363 LeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlnThrHis 382
 Db 2641 TTAAGAAGCAACAGATCCAGGACCAATATGAACAAATGAACACAGCAAGAGTACAA 2700

QY 383 LeuLysGlnGlnLeuAlaLeuTyrThrLysPheGluGluPheGlnAsnThrLeuSer 402
 Db 2701 CTAATAACAGCAGCTTTCTTTATATGATAAATTTGAAGAAATTCAGACTTACTATGGCA 2760
 QY 403 LysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysIle 422
 Db 2761 AAAAGCAATGAACCTTTTACACCTTCAGCAGGAAATGGAAGATGACAAAGAAAT 2820
 QY 423 LysLysLeuGluLysGluThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAla 442
 Db 2821 AAAAATCTGAAAGAAACAAATAATATGCGTACCAATGGGAAACATATAAAGCA 2880
 QY 443 LeuLeuGluMetAlaGluLysThrValArgAspLysGluGluGluLysVal 462
 Db 2881 CTTCTGAGATGGCCGAGAGAAACTGCTCGTATAAAGAGTACAGGCTTTTCAATA 2940
 QY 463 LysIleGlnArgLeuLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsn 482
 Db 2941 AAATGGAACGTTAGAGAGCTGTGAGGCTCTTCAGACAGAGAAATGAGCTCAAC 3000
 QY 483 LysArgValGlnAspLeu 488
 Db 3001 GAGAAGCTCGAAGTCTCTG 3018
 RESULT 25
 LOCUS CA324134 778 bp mRNA linear EST 09-JUL-2003
 DEFINITION UI-M-FY0-cco-f-03-0-UT.r1 NIH BMAP_FY0 Mus musculus cDNA clone
 IMAGE:6822244 5', mRNA sequence.
 ACCESSION CA324134
 VERSION CA324134.1 GI:24542232
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 778)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-f@mail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.
 Location/Qualifiers
 1..778
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:682244"
 /tissue_type="whole brain"
 /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH BMAP FY0"
 /notes="Organ: Brain; Vector: pYX-Asc; Site 1: Ecor I;
 Site 2: Not I; The library was constructed according
 Benaldo, Lennon and Soares, Genome Research, 6:791-806.
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with Ecor I adaptor, digested with NotI and then cloned
 directionally into pYX-Asc vector. The library tag

FEATURES
source

sequence located between the Not I site and the polyA tail is ACCGAGACAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:	2,99e-63	Length:	778
Pred. No.:	1136.00	Matches:	233
Score:	89.77%	Conservative:	4
Percent Similarity:	88.26%	Mismatches:	16
Best Local Similarity:	40.79%	Indels:	11
Query Match:	14	Gaps:	3
DB:			
US-10-023-523-44 (1-546) x CA324134 (1-778)			
QY	33	ArgProSerGlnAlaAlaProAlaValGluAlaGluGlyProGlySer	----- 48
Db	3	CGACCCAGACAGACAGCTCTCT	
QY	49	SerGlnAlaPro-ArgLysProGluGlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLe	68
Db	48	AGCCAGGCTCTCTGGGAGACAGCGGGGCTCGAGCTAAAGCAGCTCAGCTGGGGCCT	107
QY	68	uArgAspValSerGluGluLeuSerArgGlnLeuGlnAspLeuSerThrTyrCysVa	88
Db	108	CTGTGACGCTCTCTAGGAGCTGAGCGGCAGTTGAAGACATCTCTGATACATCTGTGT	167
QY	88	lAspAsnAngGlnGlyProGlyCgluAspGlyAlaGlnGlyGluProAlaGluProGl	108
Db	168	GGACAAATATGAGGAGGCGCCGCTGAGGAGGAGCAGAGGTGAGCCACTGAGCCGGA	227
QY	108	uAspAlaGluLysSerArgThrTyrValAlaAlaArgAsnGlyGluProGluPro---	127
Db	228	AGACCGGAGAGTCCCGAACCTATGACCCAGGAATGGGAGCGCTGAACAGGCAATCC	287
QY	127	oValValAsnGlyGluLysGluProSerLysGlyAspProAsnThrGluGluLeuArgGl	147
Db	288	AGTCGTCAACGGCGAAGAGACCTCTTAAGGAGAGCTCTGAACAGAGAGATCCGAGC	347
QY	147	nSerAspGluValGlyAspArgAspHisArgArgProGlnGluLysLysAlaLysGl	167
Db	348	GAGTGATGAGTTGGAGACCGAGACCATCGAGGCCACAGGAGAGAGAAAGCAAGGG	407
QY	167	yLeuGlyLysGluLysThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGl	187
Db	408	TCTAGGGAAGGAGATCACTCTGCTGATGACAGACTGAACACGCTGAGTACCCAGAGA	467
QY	187	uLysLeuAlaAlaLeuCysLysLysValAlaGluLeuLeuGluGluHisArgAsnSerGl	207
Db	468	GAAGCTGGCTGCTGCTGACAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	527
QY	207	nLysGlnMetLysLeuGlnLysGlnSerGlnLeuValGlnGluLysAspHisLe	227
Db	528	GAAGCAGATGAAGCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	587
QY	227	uArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGl	247
Db	588	GCAGGGGGAACACAGAGAGCTCTCTGCGCCGAGAGAGAGAGAGAGAGAGAGAGAG	647
QY	247	uLeuGlnArgHisAsnArgSerLeuLysGluGlyValGlnArgAlaArgGluGluGl	267
Db	648	GCTGCAACGCAACACCGGTCCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	707
QY	267	uGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGl	287
Db	708	GGAGAGCGCAAGAGAGTGTCTTCACTTCCAGGTGACACTGATGATTCAGCTGCA	767
QY	287	nMetGluGln	290
Db	768	GATGGACAG	777

RESULT 26

CB244426

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CB244426

UI-M-FY0-cdg-i-09-0-UI.r1 NIH BMAP_FY0 Mus musculus cDNA clone

IMAGE:6833074 5', mRNA sequence.

CB244426

CB244426.1 GI:28366070

EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;

1 (bases 1 to 727)

NIH-MGC http://mgc.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgaabs-remail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa

cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)

Seq primer: PYX-5.

Location/Qualifiers

1. .727

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:6833074"

/tissue_type="whole brain"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NIH BMAP FY0"

/notes="Organ: Brain; Vector: PYX- Asc; Site 1: EcoR I;

Site 2: Not I; The library was constructed according

Bonaldi, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT

primer containing a Not I site. Double strand cDNA was

size selected according to mRNA size fraction, ligated

with EcoR I adaptor, digested with NotI and then cloned

directionally into PYX-Asc vector. The library tag

sequence located between the Not I site and the polyA tail

is AGCGAGACAG. This library was created for the University

Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the

Developing Mouse Nervous System', supported by National

Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,

program coordinator."

ORIGIN

Alignment Scores:

Pred. No.: 2,97e-63

Score: 1135.50

Percent Similarity: 96.28%

Best Local Similarity: 95.87%

Query Match: 40.77%

DB: 14

Length: 727

Matches: 232

Conservative: 1

Mismatches: 6

Indels: 3

Gaps: 1

US-10-023-523-44 (1-546) x CB244426 (1-727)

QY 111 GluLysSerArgThrTyrValAlaArgAsnGlyGluProGluPro---ThrProValVal 129

Db 6 GAGAGTCCCGAACCTATGCGAGCCAGGATGGGAGGCTGAACAGGCAATTCAGTCGTC 65

QY 130 AsnGlyGluLysGluProSerLysGlyAspProAsnThrGluGluLeuArgGlnSerAsp 149

Db 480 CATAACGACGCAAACTCCAACTGCGCCAGGAGAAATATGAGCTAGCCGAGAGGCTCAAG 539
 QY 311 LysLeuileGluGlnTyrGluLeuArgGluGluHisIleAspLysValPheLysHisLys 330
 Db 540 AAGTTGATCGAGCAATACGAGCTTCGTGAGGAGCATATCGCAAGATCTTCANACATAAG 599
 QY 331 AspLeuGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGlu 350
 Db 600 GACCTGACGACGACGACTAGTGGACGCCAAGCTCCAGCAGGCGCCAGGAGATGCTGAGGAG 659
 QY 351 AlaGluGluArgHisGlnArgGluLysAspPheLeuLeu 363
 Db 660 GCAGAGGAGCGCCACCGAGGAGAGGAGTTCTCTCTG 698

RESULT 28
 BX855901 780 bp mRNA linear EST 16-DEC-2003
 LOCUS AGENAE Rainbow trout multi-tissues substracted library
 DEFINITION (tcay) Oncorhynchus mykiss cDNA clone tcay0039b.j.10 Sprim, mRNA
 sequence.
 ACCESSION BX855901
 VERSION BX855901.1 GI:39952895
 KEYWORDS EST.
 SOURCE Oncorhynchus mykiss (rainbow trout)

ORGANISM
 Oncorhynchus mykiss
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 1 (bases 1 to 780)

REFERENCE
 AUTHORS Govoroun, M., Guiguen, Y. and Le Gac, P.
 TITLE Construction and primary characterization of normalized cDNA
 JOURNAL libraries in rainbow trout, Oncorhynchus mykiss
 COMMENT Unpublished (2003)
 CONTACT: Guiguen Y

INRA - SCRIBE
 Campus de Beaulieu, RENNES cedex, 35042, France
 Tel: 02.23.48.50.09
 Fax: 02.23.48.50.20
 Email: Yann.Guiguen@beaulieu.rennes.inra.fr
 Sequence cleaned of vector, adaptor and repetitions. Contact us
 at sigenasupport@jouy.inra.fr to obtain the chromatogram of this
 sequence.
 Plate: 0039 row: j column: 10
 Seq primer: M13K.

FEATURES
 SOURCE
 1..780
 Location/Qualifiers
 /organism="Oncorhynchus mykiss"
 /mol_type="mRNA"
 /db_xref="taxon:8022"
 /clone="tcay0039b.j.10"
 /tissue_type="adipose tissue, blood, brain,
 differentiating gonads, gills, interrenal, intestine,
 kidney, liver, muscle, ovary, pituitary, testis"
 /dev_stage="from embryos to adults"
 /lab_host="DH10B"
 /clone_lib="AGENAE Rainbow trout multi-tissues substracted
 library (tcay)"
 /notes="Vector: pT73D-pac; Rainbow trout multi-tissues -
 normalized + 1 subraction (tcay); Clone distribution:
 AGENAE Resource centre, Francois PIUMI,
 Francois.Piumi@jouy.inra.fr, INRA, CEA Radiobiologie et
 Etude du genome (LREG), Domaine de Vilvert, 78352,
 Jouy-en-Josas cedex, FRANCE"

ORIGIN

Alignment Scores:
 Pred. No.: 7,09e-61 Length: 780
 Score: 1099.00 Matches: 215
 Percent Similarity: 92.25% Conservative: 23
 Best Local Similarity: 83.33% Mismatches: 20
 Query Match: 39.46% Indels: 0
 DB: 13 Gaps: 0

US-10-023-523-44 (1-546) x BX855901 (1-780)

QY 170 LysGluileThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeu 189
 Db 6 GAGGAGATCACCTCGTGTGATGAGACCCCTGAACTACTCTGAGTACTCCAGAGATAAATG 65
 QY 190 AlaAlaLeuCysLysLysTyrAlaGluLeuLeuGluHisIleAsnSerGlnLysGln 209
 Db 66 ACAGGCTCTGTAAAGATATGCTGAATCTCTGGAGGCGCGGAAACACACCAAGAGCTA 125
 QY 210 MetLysLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGly 229
 Db 126 ATCCGGGCTCTGCAGAAAGAACAGTCTCAGCTGTGTGAGGAGCCAGGACCCACCTGAGGA 185
 QY 230 GluHisSerLysAlaValLeuAlaArgSerLysLeuLeuSerLeuCysArgGluLeuGln 249
 Db 186 GAACACAGCAAGACCATCTCTGCACGACGCAAACTGGAGTGGCTCTGTGGGAGCTGAG 245
 QY 250 ArgHisAsnArgSerLysLysGluGlyValGlnArgAlaArgGluGluGluLys 269
 Db 246 AGACACACACCGCACACTCAAGGATCAGCGGATGCGAGCTGCACGTGTGGAGGAGGAG 305
 QY 270 ArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuMetGlu 289
 Db 306 AGGAAGGAGTGCACGACACTTCCAGGTGACTCTGTATGACATCCAGGCCAGATGGAG 365
 QY 290 GlnHisAsnGluArgAsnSerLysLeuArgGlnGlnAsnMetGluLeuAlaGluArgLeu 309
 Db 366 CAGCATGACGAGAGGAATGCCAGTCTTAGGCTGGAGAACTCTGAGCTGGCAGAGAGCTC 425
 QY 310 LysLysLeuileGluGlnTyrGluLeuArgGluGluHisIleAspLysValPheLysHis 329
 Db 426 AAGAAACTCATCCAAACAGTACGAAAGTACGAGGAGGACACATTGACAAAGTGTCAAGCAT 485
 QY 330 LysAspLeuGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLys 349
 Db 486 AAGAACTGACGACACAGTTGTTGATGCTAAGTACACAGGCTCAGGGTCTCTCATTA 545
 QY 350 GluAlaGluGluArgHisGlnArgGluLysAspPheLeuLysGluAlaValGluSer 369
 Db 546 GAGTCAGAGGACCGCCATCACAGAGAGAGGACTTTCTGCTGAAGAGGAGCGGAGAGTCT 605
 QY 370 GlnArgMetCysGluLeuMetLysGlnGlnThrHisLysLysGlnGlnLeuAlaLeu 389
 Db 606 CAGAGATGTGTGAGCTCATGACGAGGAGGTGCTCTCAAAACACAGCTGTCTCTTA 665
 QY 390 TyrThrGluLysPheGluPheGlnAsnThrLeuSerLysSerSerGluValPheThr 409
 Db 666 TACACAGAGAAGTTTGGAGGTTCACAGCAGCTCTGTCCAGAGCAACAGAGGTGTTCCACC 725
 QY 410 ThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLysLysLysLys 427
 Db 726 ACATTCAACAGGAGATGGAGAGATCAACAAAGAGATCAAGAGGCTGGAGAG 779

RESULT 29

BQ950691

LOCUS

BQ950691

DEFINITION

AGENCOURT 8752844 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6335682

5', mRNA sequence.

ACCESSION

BQ950691

VERSION

BQ950691.1

KEYWORDS

EST.

SOURCE

Mus musculus

ORGANISM

Mus musculus

REFERENCE

1 (bases 1 to 930)

AUTHORS

NIH-MGC http://mgi.nci.nih.gov/

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

CONTACT: Robert Strausberg, Ph.D.

Email: csapsb@mail.nih.gov

Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman,

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Ph.D.
cDNA Library Preparation: ResGen, Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13798 row: h column: 19
High quality sequence stop: 560.
Location/Qualifiers
1. 930
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6335682"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 130"
/note="Organ: oocytes; Vector: pCMV-SPORT6.1; Site 1:
EcoRV; Site 2: NotI; Cloned unidirectionally. Primer:
Oligo dt. Average insert size 1.95 Kb. Constructed by
ResGen, Invitrogen Corp. Note: this is a NIH MGC Library."

FEATURES
source
Alignment Scores:
Pred. No.: 8,79e-61 Length: 930
Score: 1099.00 Matches: 232
Percent Similarity: 84.86% Conservative: 9
Best Local Similarity: 81.69% Mismatches: 36
Query Match: 39.46% Indels: 7
DB: 13 Gaps: 3

US-10-023-523-44 (1-546) x BQ950691 (1-930)

Qy 1 MetLysAsnGlnAspLysLysAsnGlyAlaLysGlnSerAsnProLysSerSerPro 20
Db 48 ATGAAGAACCAAGCAAAAGAAATGGCTGCGCAACACTCCAACTCGAAGCGGACGCCG 107
Qy 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
Db 108 GGGCAACGGGAAGCAGGAGCGGAGGAGCCCATGAGCAGCCAGACACACAGCTCTCTGG 167
Qy 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
Db 168 GCGGAAGCCGGAAGGT--TCCACAGCCAGGCTCTCTGGAGACCGAGGCGCTCGAGCT 224
Qy 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
Db 225 AAAGCAGCTCACGCTGGGGCCCTCTGTGACGCTCTGAGGAGCTGAGCGGCGAGTTGGAA 284
Qy 81 AspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAla 100
Db 285 GACATCCTGAGTACATACTGTGTGACACACATCAGGAGGCGCGGCTGAGGAGGAGCA 344
Qy 101 GlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsn 120
Db 345 CAGGGTGAGCCCACTGAGCGGAAGACACGGAAGAGTCCCGAACCTATGACGCCAGGAAT 404
Qy 121 GlyGluProGluPro--ThrProValValAsnGlyGluLysGluProSerLysGlyAsp 139
Db 405 GGGAGCCCTGAAACCGGCAATTCAGTCTGTCACGCGGAGAGAGACCTCTAAGGAGAG 464
Qy 140 ProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAspHisArgPro 159
Db 465 CTGTGAACAGAGAGAGATCCGAGCGAGTGTGAAGTTGGAGACCGAGACCATCGGAGGCA 524
Qy 160 GlnGluLysLysLysAlaLysGlyLeuGlyLysGluLeuThrLeuMetGlnThrLeu 179
Db 525 CAGGAGAAGAAAGAACCCAGGGCTTAGGGAAGGAGATCACTCTGCTGATGACACACTG 584
Qy 180 AsnThrLeuSerThrProGluGluLysLeuAlaLeuLysLysLysTyrAlaGluLeu 199
Db 585 AACACGCTGAGTACCCACAGGAGAGAGCTGGCTGCATCTGCAAGAGATGATCTGAGCTG 644

Qy 200 LeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuLysLysGlnSerGln 219
Db 645 CTGGAAGAGCATCGGAACCTCGAAGACGATGAAGCTCTCTCAGAAAGAGAGAGCCAG 704
Qy 220 LeuValGlnGlnLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSer 239
Db 705 CTGCTGAGGAGAGAGACCATCTCGGAGGAGACACAGCCAGGCTCTCTGGCCCCGAGC 764
Qy 240 LysLeuGluSerLeuLysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGluG 259
Db 765 AAGCTTGAGAGTCTGTGCGCGGAGCTGCACGGGACCCCGGTCCTCTGAAAGAGAAG 824
Qy 259 lyVal-----GlnArgAlaArgGluGluGluGluLysArgLysGluValThrSerHis 276
Db 825 GCGTGCGCGCGAGCCCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 884
Qy 277 PheGln 278
Db 885 TTCAG 890

RESULT 30
AL637771
LOCUS
DEFINITION AL637771 XGC-neurula Silurana tropicalis cDNA clone TNeu020c17 5',
mRNA sequence.
ACCESSION AL637771
VERSION AL637771.2 GI:38216372
KEYWORDS EST.
SOURCE Silurana tropicalis (western clawed frog)
ORGANISM Silurana tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Silurana.
REFERENCE 1 (bases 1 to 702)
AUTHORS Croning, M.D.R., Ashurst, J.L., Taylor, R., Zorn, A.M. and Rogers, J.
TITLE Sanger Xenopus tropicalis EST project 2001 (11_2003)
JOURNAL Unpublished (2003)
COMMENT On Nov 7, 2001 this sequence version replaced gi:16789750.
Contact: Huckle E
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
cDNA was oligo dt primed from sug of poly A+ RNA from neurula.
EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
5' end and NotI at the 3' end.
Vector: pCS107; Site1: EcoRI; Site2: NotI
Host: Escherichia coli DH10B
Sanger Xenopus tropicalis EST project 2001
TROPICALIS SEQUENCE ID: TNeu020c17.plkSP6
Sequencing primer: SP6.
Location/Qualifiers
1. 702
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clones="TNeu020c17"
/dev_stage="neurula"
/lab_host="Escherichia coli DH10B"
/clone_lib="XGC-neurula"
/notes="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
was oligo dt primed from sug of poly A+ RNA from neurula.
EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end."

ORIGIN
Alignment Scores:
Pred. No.: 8,89e-60 Length: 702
Score: 1081.00 Matches: 213
Percent Similarity: 96.14% Conservative: 11
Best Local Similarity: 91.42% Mismatches: 9
Query Match: 38.82% Indels: 0

DB: 9 Gaps: 0

US-10-023-523-44 (1-546) x AL637771 (1-702)

QY 169 GlyLysGluLeuThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLys 188
 Db 2 GGAAGGAGATCATATTATATGGCAGACTCTGCACACTCAGCACCCAGGAGAAAG 61

QY 189 LeuAlaAlaLeuCysLysLysTyrAlaGluLeuLeuGluLeuHisArgAsnSerGlnLys 208
 Db 62 CTGACTGCCCTGTGTAGAGATATGAGAACTGTGGAGGAGCAGGACATCTCAGAAAG 121

QY 209 GlnMetLysLeuLeuGlnLysGlnSerGlnLeuValGlnGluLysAspHisLeuArg 228
 Db 122 CAAATGAGGATACGCAAAAAAAGACCCAGCTTATTCAAGAAAGGATCTACTTCGT 181

QY 229 GlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLeu 248
 Db 182 AATGACGACAGCAAGCTATCTGGCAAGAGCAAGCTCGAAGTTGTGCGAGAGCTTA 241

QY 249 GlnArgHisAsnArgSerLysLysGluGluGlyValGlnArgAlaArgGluGluGlu 268
 Db 242 CAGAGCACACACCTACACTAAGGAGGAGGAGGTACAGAGGGCTCTGAGGAGGAGAG 301

QY 269 LysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuMet 288
 Db 302 AAACGCAAAAGAGTAACCTCCCATTTCCAAAGTGACGCTCAATGACATCCAGTCACAGATG 361

QY 289 GluGlnHisAsnGluArgSerLysLeuArgGlnGluAsnMetGluLeuAlaGluArg 308
 Db 362 GAGCAGCACATGAAGCTAATGTAACCTCGGTCAAGAGAAATGTGAGCTTCTGATCA 421

QY 309 LeuLysLysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAspLysValPheLys 328
 Db 422 CTTAAGAAGCTTATCGAGCAATATAGCTAAGAGAGAGACATTCACAAAGTTTCAAA 481

QY 329 HisLysAspLeuGlnGlnGlnLeuValAlaLysLeuGlnGlnAlaGlnLeuMetLeu 348
 Db 482 CATAAGGATCTTCAACAGCAACTGTGGATGCAAGAGTTTCAGACAGCAAGAAATGTG 541

QY 349 LysGluAlaGluGluArgHisGlnArgGluLysAspPheLeuLysLysAlaValGlu 368
 Db 542 AAAGAAGTGGAGAGAGCTCACACGGGGAAGAGGTTCTGTTGAAGAGGACATCGAA 601

QY 369 SerGlnArgMetCysGluLeuMetLysGlnGlnGlnThrHisLeuLysGlnLeuAla 388
 Db 602 TCCAGCGCATGTGTGAACATAATGAACAGCAGAGAGACCCACCTCAAAACAGTAGCA 661

QY 389 LeuTyrThrGluLysPheGluGluPheGlnAsnThrLeu 401
 Db 662 CTGTATACAGAGAGTTTGAGGAGTTTCAAAACACCTTG 700

RESULT 31
 BM963869
 LOCUS
 DEFINITION UI-M-EQO-bwt-b-08-0-UI.r1 NIH_BMAP_EQO Mus musculus cDNA clone
 IMAGE:5698471 5', mRNA sequence.

ACCESSION BM963869
 VERSION
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 JOURNAL 1 (bases 1 to 735)
 COMMENT NIH-MGC <http://mgi.nhl.nih.gov/>.
 Tissue Procurement: Dr. James Lin, University of Iowa
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.

FEATURES

source

Location/Qualifiers

1..735
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:5698471"
 /tissue_type="whole brain"
 /dev_stage="embryo 15.5 dpc"
 /lab_host="DHIOB (T1 phage resistant)"
 /clone_lib="NIH_BMAP EQO"
 /note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was size selected according to mRNA size fraction,
 ligated with EcoR I adaptor, digested with Not I, and then
 cloned directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA
 tail, is CAGCCAGCAG. This library was created for the
 University of Iowa Mouse Brain Molecular Anatomy Project
 (BMAP). Gene Discovery in the Developing Mouse Nervous
 System", supported by National Institutes of Mental Health
 (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:
 Pred. No.: 4,12e-59 Length: 735
 Score: 1071.00 Matches: 214
 Percent Similarity: 89.02% Conservative: 5
 Best Local Similarity: 86.99% Mismatches: 25
 Query Match: 38.46% Indels: 2
 DB: 12 Gaps: 2

US-10-023-523-44 (1-546) x BM963869 (1-735)

QY 9 GlyAlaAlaLysGlnSerAsnProLysSerProGlyGlnProGluAlaGlyProGlu 28
 Db 1 GGGCTCCCAACACTCCAACTCGAAGGGCAGCCCGGGCAACGGGAAGCAGACCGGAG 60

QY 29 GlyAlaGlnGluArgProSerGlnAlaAlaProAlaValGluAlaGluGlyProGlySer 48
 Db 61 GGAGCCCATCGACACCCACAGACAGCTCTCTGGGGCGGAGGCCAAGT---TCACA 117

QY 49 SerGlnAlaProArgLysProGlyGluAspGlyValaGlnAlaArgThrAlaGlnSerGlyAlaLeu 68
 Db 118 AGCCAGGCTCTCTGGGAAGACCGAGGGGGCTCGAGCTAAAGCAGCTCAGCCTGGGGCCCTC 177

QY 69 ArgAspValSerGluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysVal 88
 Db 178 TGTGACGCTCTCTGAGGAGCTGAGCCGGCAGTTGGAAGACATCTCTAGTACATCTGTGTG 237

QY 89 AspAsnAsnGlnGlyGlyProGlyGluAspGlyValaGlnGlyGluProAlaGluProGlu 108
 Db 238 GACAAACATCAGGAGGCGCCGGCTGAGGAGGGAGCAGAGGCTGAGCCCTGAGCCGGA 297

QY 109 AspAlaGluLysSerArgThrTyrValAlaAlaArgAsnGlyGluProGluPro---ThrPro 127
 Db 298 GACACGAGAGAGTCCCAACCTATGAGCCAGCAAGATGGGGAGCCTGAACACGAGCATCCA 357

QY 128 ValValAsnGlyGluLysGluProSerLysGlyAspProAsnThrGluGluLeuArgGln 147
 Db 358 GTGCTCAACGCGGAGAGGAGACCTCTAAGGGAGAGCCTTGAACAGAGAGATCCGAGCG 417


```

source
1. 861
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="FVB/N"
/clone="IMAGE:5043572"
/lab_host="NIH CGAP Kid14"
/clone_lib="NIH CGAP Kid14"
/notes="Organ: kidney; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.75 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library. |"

ORIGIN

Alignment Scores:
Pred. No.: 2,19e-58 Length: 861
Score: 1061.00 Matches: 220
Percent Similarity: 88.63% Conservative: 6
Best Local Similarity: 86.27% Mismatches: 12
Query Match: 38.10% Indels: 17
DB: 12 Gaps: 3

US-10-023-523-44 (1-546) x B1102887 (1-861)
QY 304 GluLeuAlaGluArgLeuLysLysLeuLeuGluGlnTyrGluLeuArgGluGluHisile 323
Db 2 GAGCTAGCCGAGAGGCTCAAGAAGTTGATCGACGATAGCAGCTTCGTGAGAGGATATC 61
QY 324 AspLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGln 343
Db 62 GACAAAGTCTTCAAAATAGGACCTGCAGCAGCAGCTAGTGCAGCCCAAGCTCCAGCAG 121
QY 344 AlaGlnGluMetLeuLysGluAlaGluArgHisGlnArgGluLysAspPheLeuLeu 363
Db 122 GCCCAGAGATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 181
QY 364 LysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlnThrHisLeu 383
Db 182 AAGGAAGCGGTGAGTCCAGAGAGTGTGCGAGCTGATGAAGCAGCAGGAGCACCACCTA 241
QY 384 LysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLys 403
Db 242 AAGCAGCAGCTGCGCTGTACACGAGAGAGTGTGAGGAGTTCAGAACACATTTCCAAA 301
QY 404 SerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysileLys 423
Db 302 ACCAGTGAAGTGTTCACCACTTCAACAGGAGATGGAACAGATGACAAAGATCAA 361
QY 423 sLysLeuGluLysGluThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLe 443
Db 362 GAAGCTGGAGAAGAAACCAACCATGTATGATCCAGTGGGAAGCAGCAACAGGCTCT 421
QY 443 uLeuGluMetAlaGluGluLysThrValArgAspLysGluLeuGluGluLysValLys 463
Db 422 GTTGAGATGCTGAAGAGAAACCGTCCGGACAAAGAGCTGGAGGGGCTCGAGTGA 481
QY 463 sIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLys 483
Db 482 AATCCAGCGCTGGAGAGCTGTGCGAGCAGCTGCAGACTGAGCGCAATGACCTCAACAA 541
QY 483 sArgValGlnAspLeuSerAlaGlyGlnGlySerLeuThrAspSerGlyProGluArg 503
Db 542 GAGGGTACAGGACCTGACTGCGAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 592
QY 503 gArgPro-----GluGlyProGlyValaGlnAl 512
Db 593 GAGGCCAGAGGCCACCACTGCTCCAGGAACAGAGGGTTGAAGCTCTGGGGCTCAACC 652
QY 512 aProSerSerProArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAl 532
Db 653 AGCCAGCTCTCCAGGGGCCACAGCGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 712
QY 532 aSerGlyGlnThrGlyProGlnGlu-----ProThrSerAla 544

```

```

Db 713 AGCAGCCAGACAGCGCTCGAGGAAGCAACCCCTGTCACTGCC 755
:::|||||
RESULT 35
CD578468 647 bp mRNA linear EST 09-JUL-2003
LOCUS UI-M-FY0-cfs-n-14-0-UI-r1 NIH_BMAP_FY0 Mus musculus CDNA clone
DEFINITION IMAGE:6855471 5', mRNA sequence.
ACCESSION CD578468
VERSION CD578468.1 GI:31742859
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (Bases 1 to 647)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mouserfi.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: PYX-5.
Location/Qualifiers
1. 647
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="CS7BL/6"
/clone="IMAGE:6855471"
/tissue_type="whole brain"
/dev_stage="embryo 13.5, 14.5, 16.5, 17.5dpc"
/lab_host="DH10B (TI phage resistant)"
/clone_lib="NIH_BMAP_FY0"
/notes="Organ: Brain; Vector: PYX- Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into PYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCGAGACAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."
ORIGIN
Alignment Scores:
Pred. No.: 1,79e-58 Length: 647
Score: 1060.00 Matches: 210
Percent Similarity: 99.53% Conservative: 2
Best Local Similarity: 98.59% Mismatches: 1
Query Match: 38.06% Indels: 0
DB: 14 Gaps: 0
US-10-023-523-44 (1-546) x CD578468 (1-647)
QY 151 ValGlyAspArgAspHisArgArgProGlnGluLysLysLysLysLysGlyLeuGlyLys 170
Db 9 ATTGGAGACCGAGACCATCGAGGCCACAGAGAGAGAGAGCAAGGCTCAGGGAAG 68
QY 171 GluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAla 190

```

```

Db      69  GAGATCACTCTGCTGATGACACACTGACACGCTGAGTACCCAGAGGAGAGCTGGCT 128
Qy      191 AlaLeuCyLysLysTyrAlaGluLeuLeuGluHisArgAsnSerGlnLysGlnMet 210
Db      129 GCACTGTGCAAGAAGATGCTGAGCTGCTGGAAGACATCGGAATCGCAGAGCAGATG 188
Qy      211 LysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGlnLysAspHisLeuArgGlyGlu 230
Db      189 AAGCTCTCTGAGAGAGAGCAGAGCGAGCTGCTGCAAGAGAGAGGACATCTGCGAGGGAA 248
Qy      231 HisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCyArgGluLeuGlnArg 250
Db      249 CACAGCAAGGCTGCTGCGCCGAGCAAGCTGTGAGAGTCTGTGCGGAGCTGCAAGCG 308
Qy      251 HisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGluGluGluLysArg 270
Db      309 CACAACCGGTCTCTGAAGAGAGAGGCGTGCAGCGAGCGCTGAGGAGGAGAGAGCGC 368
Qy      271 LysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuMetGluGln 290
Db      369 AAGAAGTCACTTCAACATTCACATTCACAGGTGACACTGAATGACATTCAGCTGCAAGTGAACAG 428
Qy      291 HisAsnGluArgAsnSerLysLeuArgGlnGlnGluAsnMetGluLeuAlaGluArgLeuLys 310
Db      429 CATACGAGCGCAAACTCCAGCTGCGCCAGAGGAGATATGAGCTAGCGAGAGGCTCAAG 488
Qy      311 LysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAspLysValPheLysHisLys 330
Db      489 AAGTTGATCAGCAATACAGCTTCGTGAGGAGCATATCGACAAAGTCTTCANACATAG 548
Qy      331 AspLeuGlnGlnGlnLeuValAlaLysLeuGlnGlnAlaGlnGlnMetLeuLysGlu 350
Db      549 GACCTGAGCAGCAGCTAGTGGACGCGCAAGCTCCAGCAGCGCCAGCAGAGATGCTGAGGAG 608
Qy      351 AlaGluGluArgHisGlnArgGluLysAspPheLeuLeu 363
Db      609 GCAGAGGAGCGGCACAGCAGAGAGAGGAGTTTCTCTG 647

RESULT 36
BU613886      653 bp mRNA linear EST 20-FEB-2003
LOCUS      UI-M-EWO-cba-p-12-0-UI.r1 NIH EMAP_EWO Mus musculus cDNA clone
DEFINITION      UI-M-EWO-cba-p-12-0-UI 5', mRNA sequence.
ACCESSION      BU613886
VERSION        BU613886.1 GI:23280101
KEYWORDS       Esi.
SOURCE         Mus musculus (house mouse)
ORGANISM       Mus musculus
REFERENCE      1. (bases 1 to 653)
AUTHORS        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
JOURNAL        National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT        NIH-MGC http://mgc.nci.nih.gov/
                Unpublished (1999)
                Contact: Robert Strausberg, Ph.D.
                Email: cgapbs-r@mail.nih.gov
                Tissue Procurement: Dr. James Lin, University of Iowa
                cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
                cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
                DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
                Clone Distribution: Clone distribution information can be obtained
                from Dr. M. Bento Soares, bent-soares@uiowa.edu
                This clone was contributed by the Brain Molecular Anatomy Project
                (BMAP)
Seq primer: pyx-5.
Location/Qualifiers
1..653
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="UI-M-EWO-cba-p-12-0-UI"

```

```

/tissue_type="whole brain"
/dev_stage="embryo 15.5 dpc"
/lab_host="DHI0B (T1 phage resistant)"
/clone_lib="NIH_EMAP_EWO"
/notes="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with EcoR I adaptor, digested with Not I, and then
cloned directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA
tail, is GTGGGTGGA. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP). Gene discovery in the Developing Mouse Nervous
System", supported by National Institutes of Mental Health
(NIMH), Hemin Chin, Ph.D., program coordinator."
ORIGIN
Alignment Scores:
Pred. No.: 2,1e-58 Length: 653
Score: 1059.00 Matches: 212
Percent Similarity: 99.54% Conservative: 3
Best Local Similarity: 98.15% Mismatches: 0
Query Match: 38.03% Indels: 1
DB: 13 Gaps: 0
US-10-023-523-44 (1-546) x BU613886 (1-653)
Qy 148 SerAspGluValGlyAspArgAspHis-ArgArgProGlnGluLysLysLysLysG1 167
Db 3 ACTCATGAAGTTGAGACCCAGACCATCCGAGGCCACAGAGAGAGAGAGCAAGGG 62
Qy 167 YLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluG1 187
Db 63 TCTAGGAGAGGAGATCACTCTGCTGATGCAGACACTGAACACACTGATGATCCCGAGGA 122
Qy 187 uLysLeuAlaAlaLeuCyLysLysTyrAlaGluLeuLeuGluGluHisArgAsnSerG1 207
Db 123 GAAGCTGGCTGCACTGTGCAAGAGTATGCTGAGCTGCTGGAAGAGCATCGGAACCGCA 182
Qy 207 nLysGlnMetLysLeuLeuGlnLysGlnSerGlnLeuValGlnGlnLysAspHisL1e 227
Db 183 GAAGCAGATGAAGCTCTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 242
Qy 227 uArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCyArgG1 247
Db 243 CCGAGGGGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 302
Qy 247 uLeuGlnArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGluG1 267
Db 303 GCTGCAACGGCACAAACCGGTCCTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 362
Qy 267 uGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuG1 287
Db 363 GGAGAGCGCCAAAGAGTCACTTCACTTCCAGTGCAGCTGAATGATCATTCAGCTGCA 422
Qy 287 nMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGlnGlnMetGluLeuAlaG1 307
Db 423 GATGAAACAGCATTAACAGAGCGAAACTCCAAAGTCTGCGCCAGAGAGAGAGAGAGAG 482
Qy 307 uArgLeuLysLysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAspLysValPh 327
Db 483 GAGGCTCAAGAGATTGATCGAGCAATACGAGCTTCTGAGGAGAGAGATATCCAAAGTCTT 542
Qy 327 eLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMe 347
Db 543 CAACATTAAGGACCTGCGAGCAGCAGCTAGTGGAGCGCCCAAGCTCCAGAGAGAGAGAG 602
Qy 347 tLeuLysGluAlaGluGluArgHisGlnArgGluLysAspPheLeu 362

```


http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.

FEATURES

Location/Qualifiers

```
1..784
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6837021"
/tissue_type="whole brain"
/dev_stage="embryo 12.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP F10"
/note="Organ: Brain; Vector: pYX-Asc; Site: 1: EcoR I; Site 2: Not I; The library was constructed according to Bernaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CAGCCACGAC. This library was created for the University of Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."
```

ORIGIN

Alignment Scores:

```
Pred. No.: 1,79e-57 Length: 784
Score: 1046.00 Matches: 215
Percent Similarity: 87.65% Conservative: 5
Best Local Similarity: 85.66% Mismatches: 29
Query Match: 37.56% Indels: 3
DB: 14 Gaps: 2
```

US-10-023-523-44 (1-546) x CB247697 (1-784)

```
QY 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerPro 20
DB 35 ATGAGAACCAACAGCAAAAGAAATGGCGCTGCCAACAACCTCCAACTCGAAGGCGAGCCG 94
QY 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
DB 95 GGCAACCGGAACAGAGCCGAGGAGCCATGGACGACCCAGACAGACAGACACTCTGGG 154
QY 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
DB 155 GCGAGAGCCGAAGT---TCCACAGCCAGGCTCTCTGGAGACCCGAGCGGCTCGAGCT 211
QY 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
DB 212 AAAGCAGCTCAGCGCTGGGGCCCTCTGTGACGCTCTGTGAGAGCTGAGCGCGCAGTTGGAA 271
QY 81 AspIleLeuSerThrTyrcysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAla 100
DB 272 GACATCTGTAGTACATCTGTGTGGACACACATCAGGAGGCGCCGCTCAGGAGGAGCA 331
QY 101 GlnGlyGluProAlaGluProGluAlaGluLysSerArgThrTyrrValAlaArgAsn 120
DB 332 CAGGGTGAGCCCACTCAGCCGGAAGACACGAGAGACTCCGAACTATCGACCCAGGAAT 391
QY 121 GlyGluProGluPro---ThrProValAlaGlnGlyGluLysGluProSerLysGlyAsp 139
DB 392 GGGAGGCTGAACAGGCAATTCACGTCGTCACCGGAGAGAGAGACCTCTAAGGAGAG 451
QY 140 ProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAspHisArgPro 159
DB 452 CCTGGAACAGAGAGATCCGAGCGAGTGATGAAGTTGGAGACCCGAGACCATCGGAGCCA 511
```

```
QY 160 GlnGluLysLysLysAlaLysGlyLeuGlyLysGluLeuThrLeuLeuMetGlnThrLeu 179
DB 512 CAGGAGAAGAAAGAACCAAGGCTCTAGGAAGAGATCACTCTGCTGATGAGACACTG 571
QY 180 AsnThrLeuSerThrProGluGluLysLeuAlaLeuAlaLeuCysLysLysTyrrAlaGluLeu 199
DB 572 AACACGCTGAGTACCCACAGAGAGAGAGCTGGCTGCCTGTGCAAGAGATATGCTGAGCTG 631
QY 200 LeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGln 219
DB 632 CTGAGAGAGATCGAAGCTCGCAAGACAGATGAGCTCTGAGAGAGAGAGAGAGAGAG 691
QY 220 LeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSer 239
DB 692 CTGTGACAGAGAGAGGA-CATCTGCGAGGGAACACAGCANGGCTCTCTGCGCCGAGC 750
QY 240 LysLeuGluSerLeuCysArgGluLeuGlnArg 250
DB 751 AAGCTTGAGAGTCTGTGCCCGGAGCTGCAACGG 783
```

RESULT 39

BG975353

LOCUS

602842995F1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:4978775 5', mRNA sequence.

ACCESSION

BG975353

VERSION

BG975353.1 GI:14362990

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 796)
NIH-MGC <http://mgs.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-xemail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth Ph.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM10975 row: n column: 24

High quality sequence stop: 780.

FEATURES

source

```
1..796
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NMRI"
/db_xref="taxon:10090"
/clone="IMAGE:4978775"
/tissue_type="tumor, gross tissue"
/dev_stage="5 months"
/lab_host="DH10B"
/clone_lib="NCI CGAP Mam4"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Priscilla Furth, NIH Reference for transgenic model: Li et al., Cell Growth and Differentiation 7, 3-11 (1996)."
```

ORIGIN

Alignment Scores:

```
Pred. No.: 3,54e-57 Length: 796
Score: 1041.50 Matches: 213
Percent Similarity: 86.61% Conservative: 7
Best Local Similarity: 83.86% Mismatches: 30
```

Query Match: 37.40% Indels: 4
DB: 12 Gaps: 3
US-10-023-523-44 (1-546) x BG975353 (1-796)

QY 1 MetLysAsnGlnAspLysAsnGlyAlaLysGlnSerAsnProLysSerPro 20
DB 30 ATGAAGAACCAAGACAAAGAAATGGCGCTGCCAACTCCAACTCGAAGGCGAGCCG 89
QY 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
DB 90 GGGCAACGGAG 149
QY 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
DB 150 GCGAAGCCGAGGT---TCCACAGCCAGGCTCTCTGGGAAGACCGAGGGGCTCGAGCT 206
QY 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
DB 207 AAAGCAGCTCAGCCTGGGGGCTCTGTGACCTCTCTGAGGAGCTGAGCGGCGAGTTGAA 266
QY 81 AspileLeuSerThrTyrcysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAla 100
DB 267 GACATCTCTGAGTACATCTGTGTGACACATCATCAGGAGAGAGAGAGAGAGAGAG 326
QY 101 GlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrcysValAlaArgAsn 120
DB 327 CAGGCTGAGCCGACCTGAGCGGAGAGACAGGAGAGAGAGAGAGAGAGAGAGAGAG 386
QY 121 GlyGluProGluPro---ThrProValValAsnGlyGluLysGluProSerLysGlyAsp 139
DB 387 GGGAGAGCTGAGCAGGAGGATTCAGCTCTCAACCGGAGAGAGAGAGAGAGAGAGAG 446
QY 140 ProAsnThrGluGluAlaArgGlnSerAspGluValGlyAspArgAspHisArgArgPro 159
DB 447 CCGTGAACAGAGGAGATCCGAGCAAGTGTGAGGTGGAGACCGAGACCATCGAGGCCA 506
QY 160 GlnGluLysLysLysAlaLysGlyLeuGlyLysGluLysGluLysLeuMetGlnThrLeu 179
DB 507 CAGGAG 566
QY 180 AsnThrLeuSerThrProGluGluLysLeuAlaLeuLysLysLysLysLysLysLys 199
DB 567 AACAGCTGTAGTACCCAG 626
QY 200 LeuGluGluHisArgAsnSerGlnLysGlnMetLysLeu---LeuGlnLysLysGlnSer 218
DB 627 CTGGAAGAGCATCGGAATCTCGCAGAGCAGCATGAAGCTCTCTGAGAGAGAGAGAG 686
QY 219 GlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArg 238
DB 687 CAGCTGTGTGAGGAG 746
QY 238 gSerLysLeuGluSerLeuGluGluGluGluGluGluGluGluGluGluGluGluGlu 251
DB 747 AAGCAAGCTTGAGAGTCTGTGCGGGAGCTGCAAGGGGAC 786

RESULT 40
AL887857
LOCUS
DEFINITION AL887857 XGC-egg Silurana tropicalis cDNA clone TEG9080c08 5', mRNA
sequence.
ACCESSION AL887857
VERSION AL887857.2 GI:38689228
KEYWORDS EST.
SOURCE Silurana tropicalis (western clawed frog)
ORGANISM Silurana tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Silurana.
1 (bases 1 to 665)
Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
Sanger Xenopus tropicalis EST project 2001 (11_2003)

Unpublished (2003)
On Sep 16, 2002 this sequence version replaced gi:22938408.
Contact: Taylor R
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: tropesanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS SEQUENCE ID: TEG9080c08.plkSP6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
cDNA was oligo dt primed from Sug of poly A+ RNA from egg.
EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
5' end and NotI at the 3' end.
Vector: pCS107; Site_1: EcoRI; Site_2: NotI
Host: Escherichia coli XL1-blue.
Location/Qualifiers
1.665
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TEG9080c08"
/dev_stage="egg"
/lab_host="Escherichia coli XL1-blue"
/clone_lib="XGC-egg"
/notes="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
was oligo dt primed from Sug of poly A+ RNA from egg.
EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end"

ORIGIN
Alignment Scores:
Pred. No.: 3,54e-57 Length: 665
Score: 1040.00 Matches: 205
Percent Similarity: 97.29% Conservative: 10
Best Local Similarity: 92.76% Mismatches: 6
Query Match: 37.34% Indels: 0
DB: 9 Gaps: 0
US-10-023-523-44 (1-546) x AL887857 (1-665)

QY 168 LeuGlyLysGluLeuThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGlu 187
DB 1 CTGGGAAGAGAGATCATATTATTATGACAGACTCTGACACACTCAGACCCAGAGAA 60
QY 188 LysLeuAlaLeuLysLysLysTyrcysAlaGluLeuGluGluHisArgAsnSerGln 207
DB 61 AAGCTGACTGCCCTGTGTAAAGTATGCAAGTCTTGGAGGAGCAGGACATCTCAG 120
QY 208 LysGlnMetLysLeuLysGlnSerGlnLeuValGlnGluLysAspHisLeu 227
DB 121 AAGCAATGAGGATCTGCAAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 228 ArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuGluGlu 247
DB 181 CGTAATGAGCAGCAGCAAGGCTATCTCTGGCAAGCAAGCAAGCTTGTGACAGAG 240
QY 248 LeuGlnArgHisAsnArgSerLeuLysGluGlyValGlnArgAlaArgGluGlu 267
DB 241 CTACAGAGACACACACCGTACACTAAAGAGGAGGAGTACAGAGGCTCTGAGGAGAA 300
QY 268 GluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspileGlnLeuGln 287
DB 301 GAGAAAGCAGCAAGAGAGTAACTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 360
QY 288 MetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGlnAsnMetGluLeuAlaGlu 307
DB 361 ATGAGCAGCAGCAATGAAGCTAATGCTAACTCGCTCAAGAGAGATGTTGAGCTTGTGAT 420
QY 308 ArgLeuLysLysLeuLeuGluGlnTyrcysAlaGluGluGluHisLysLysValPhe 327
DB 421 CGACTTAAGAGCTTATCGAGCAATGAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAG 480

```

QY 328 LysHisLysAspLeuGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMet 347
Db 481 AAACATAGGATCTTCACAGCAACTGTGGATGCGAGCTTCAGCAGCACAAGAATG 540
QY 348 LeuLysGluAlaGluGluArgHisGlnArgGluLysAspPheLeuLeuLysGluAlaVal 367
Db 541 CTGAAGAAGTGAAGAGCGTCACAGCGGGAAGAGAGTTCGTGTGAAAGAGCGCAGTC 600
QY 368 GluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLysGlnGlnLeu 387
Db 601 GAATCCCGCGCATGTGTGAACATAAGAAACAGCAAGAGACCCACCTCACAACACAGTTA 660
QY 388 Ala 388
Db 661 GCA 663

RESULT 41
BU231392
LOCUS 603947670F1 CSEQCHN23 Gallus gallus cDNA clone ChEST902117 5', mRNA
DEFINITION
ACCESSION BU231392
VERSION BU231392.1 GI:25473103
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
REFERENCE 1 (bases 1 to 782)
AUTHORS Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.O.
TITLE A Comprehensive Collection of Chicken cDNAs
JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE 22335534
PUBMED 12445392
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1. .782
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="ChEST902117"
/dev_stage="22"
/lab_host="DH10B"
/clone_lib="CSEQCHN23"
/note="Organ: heads; Vector: pBluescript II KS(+); Site_1:
EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following first strand reaction, double-stranded cDNA
was bluntended, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al. PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

ORIGIN
Alignment Scores: 6.73e-57 Length: 782
Pred. No.: 1037.00 Matches: 193
Score:

```

```

Percent Similarity: 91.51% Conservative: 44
Best Local Similarity: 74.52% Mismatches: 22
Query Match: 37.24% Indels: 0
Db: 13 Gaps: 0

US-10-023-523-44 (1-546) x BU231392 (1-782)

QY 215 LysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGluGluHisSerLysAla 234
Db 5 AAAAGCAAGCAACAAGTGTGAAGGAGAAAGTCCACTTCAGAGCGAGCATAGCAAGGCC 64
QY 235 ValLeuAlaArgSerLysLeuGluSerLysCysArgGluLeuGlnArgHisAsnArgSer 254
Db 65 ATTTCGCGACGCGACAACTCGAATCTCTCTGCGGAGCTTCAGCGTCTATACAGACT 124
QY 255 LeuLysGluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThr 274
Db 125 TTAAGGAGGAGAAACATCCAGCAAGCACGTGAAGAGGAGGAGGAGGAGGAGGAGCAACT 184
QY 275 SerHisPheGlnValThrLeuAsnAspLleGlnLeuGlnMetGluGlnHisGlnArg 294
Db 185 GCACACTTCAGATCACTCTGAATGAATTCAGGCTCAACTGGAGCAGCAGCATATACAT 244
QY 295 AsnSerLysLeuArgGlnGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuGlu 314
Db 245 AACGCCAAGCTCGCTCAAGAAATATTGAGCTGGGGGAGGAGGAGGAGGAGGAGGAGGAG 304
QY 315 GlnTyrgluLeuArgGluGluHisLysValPheLysValPheLysHisLysAspLeuGln 334
Db 305 CAGTATGCTTCGCGAGAGAGCATATTGATAAGTGTTCAGCATTAAGCAAGCAAGTTCAG 364
QY 335 GlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluGluArg 354
Db 365 CAACTGTGGATGCCAACTCCAGCAAACTCCAGCAAACTATTAAGCAAGCAAGGAGGAG 424
QY 355 HisGlnArgGluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGlu 374
Db 425 CATCAGCGGGAGAGAGAGTTCCTGCTAAAGAGAGCTACTGAGTCCAGACACAAATGTGAA 484
QY 375 LeuMetLysGlnGlnGluThrHisLysGlnGlnLeuAlaLeuTyrrThrGluLysPhe 394
Db 485 CAGATGAAGCAACAGGAGACACAGCTGGAAGCAGCAGCTTCTCTTACATGATTAATTT 544
QY 395 GluGluPheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGlu 414
Db 545 GAGGAATTCAGACTACCATGGCCAAAGTAATGAAGTGTTCACAACTTCAGGCAAGAA 604
QY 415 MetGluLysMetThrLysLysLysLysLysLysLysLysLysLysLysLysLysLys 434
Db 605 ATGAGAGAGATGACTAAGAGAGATCAAGAACTGGAAGGAGGAGGAGGAGGAGGAGGAG 664
QY 435 ArgThrGluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAsp 454
Db 665 AATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 724
QY 455 LysGluLeuGluGlyLeuGlnValLysLysLysLysLysLysLysLysLysLysLys 473
Db 725 AAAGAATACAAAGGCTTCAGATAAAAGTGGAGCGTTGGAGAGGAGTTCAGGAGGCC 781

RESULT 42
CA793605
LOCUS AGENCOURT_11044319 NICHD_XGC_Emb1 Xenopus laevis cDNA clone
DEFINITION IMAGE:6855533 5', mRNA sequence.
ACCESSION CA793605
VERSION CA793605.1 GI:26042492
KEYWORDS EST.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 862)

```

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>,
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps@mail.nih.gov
Tissue Procurement: Martha Rebert, Steven L. Klein, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
plate: LLAM14488 row: a column: 20
High quality sequence stop: 687.
Location/Qualifiers
1. .862
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:86553"
/tissue_type="embryo (stage 10)"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHD XGC Embl"
/notes="vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI;
Cloned unidirectionally. Primer: Oligo dt. Average insert
size 1.55 kb. Constructed by Life Technologies. Note: This
is a Xenopus Gene Collection (XGC) library."

ORIGIN

Alignment Scores:

Pred. No.: 7,596-57 Length: 862
Score: 1037.00 Matches: 198
Percent Similarity: 88.85% Conservative: 49
Best Local Similarity: 71.22% Mismatches: 31
Query Match: 37.24% Indels: 0
DB: 14 Gaps: 0

US-10-023-523-44 (1-546) x CA793605 (1-862)

QY 202 GluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuVal 221
DB 2 GAGCTAGGAATTCAAAACAGATGAGGTTTTCAGAGACGCGAGTCTCAAGTGTA 61
QY 222 GlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeu 241
DB 62 AAAGAAAGGTTTCATCTCGAGTGAGACAGACAGAGCGCATCTGCAGGAAGCAACTT 121
QY 242 GluSerLeuLysArgGlnLeuGlnArgHisAsnArgSerLysLysGlnGluGlyValGln 261
DB 122 GAATCTCTCGCGTGAGCTTCAGAGACAAATAAACTTTAAGAGAAATCTTCAT 181
QY 262 ArgAlaArgGluGluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeu 281
DB 182 CAAGCAGGGAATPACGAGAGAGCGCAAGGAGCCACTGCTCACITTCAGTGCATTG 241
QY 282 AsnAspLeuGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGlu 301
DB 242 AATGACATACAGCAGATGAGGACCAATGATGCTTACATGCTTAACTTCGCCAGAA 301
QY 302 AsnMetGluLeuAlaGluArgLeuLysLysLeuLeuGlnLysGlnLeuArgGluGlu 321
DB 302 AATGCTGAACTGGAGATAAATTAAGAGCTTATTGACATATGCACTGAGAGAGAG 361
QY 322 HisLysAspLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspLysLeu 341
DB 362 CACATTGACAAAGTCTTCAACATAAGGAACCTTCAACAACTATTGATCCAAACTT 421
QY 342 GlnGlnAlaGlnGlnMetLysLysGluAlaGluGluArgHisGlnArgGlnLysAspPhe 361
DB 422 CACAAACACACTGCTTATGAGGAGAGCGGAGAGAGGATCAGCGGAGAGAGGTTT 481
QY 362 LeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThr 381

DB 482 TTGTTGATGAGCACTGAAACCAAGCAAGTTTGAAGATTATACACACAGAAAGC 541
QY 382 HisLeuLysGlnGlnLeuAlaLeuThrThrGluLysPheGluGluPheGlnAsnThrLeu 401
DB 542 CAACCTGAACAGACAGCTTTCACITTTACATGATAAATTTGAGGAATTTCAACACCAATG 601
QY 402 SerLysSerSerGluValPheThrThrPheLysGlnGlnMetGluLysMetThrLysLys 421
DB 602 GCAAAAGCAATGAACTTTTACACATTTGCCAGAAATGAGAGATGACTAAGANA 661
QY 422 IleLysLysLeuGluLysGluThrThrMetTyArgSerArgTrpGluSerSerAsnLys 441
DB 662 ATAAAGAAACTTGAAAGGAAAGTGTCTCTGCCCAAAAGTGGGAAATAACACAAA 721
QY 442 AlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGluLeuGluGlyLeuGln 461
DB 722 CGCTACTGCAATGCGCTGAGAGAAACTGTACGACAGAGGAATCAAGAGCTCTCCAG 781
QY 462 ValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsn 479
DB 782 GTTAATTAGAGCGCTAGAGAAACTCNTGCGGCACCTTTCAACCGAGAGAAAC 835
BQ389733 654 bp mRNA linear EST 22-MAY-2002
NISC mq09c06.y1 NICHD XGC Emb5 Silurana tropicalis cDNA clone
IMAGE:5308018 5', mRNA sequence.
BQ389733 1 GI:21077420
EST.
Silurana tropicalis (western clawed frog)
Silurana tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Silurana.
1 (bases 1 to 654)
NIH-XGC <http://image.llnl.gov/image/html/xenopuslib.info.shtml>.
National Institute of Child Health and Human Development, National
Cancer Institute, Xenopus Gene Collection
Unpublished (2002)
Contact: Robert Strausberg, Ph.D.
Email: cgaps@mail.nih.gov
cDNA Library Preparation:
cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC)
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
plate: LLAM1780 row: E column: 11
seq primer: M13RPI reverse primer (ABI).
Location/Qualifiers
1. .654
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:5308018"
/tissue_type="gastrula"
/dev_stage="embryo, stages 10-13"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHD XGC Emb5"
/note="vector: pCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV;
Cloned unidirectionally. Primer: Oligo dt. Average insert
size 2.0 kb. Constructed by Invitrogen. Note: This is a
Xenopus Gene Collection (XGC) library."

ORIGIN

Alignment Scores:

Pred. No.: 8,42e-57 Length: 654
Score: 1034.00 Matches: 204
Percent Similarity: 98.16% Conservative: 9
Best Local Similarity: 94.01% Mismatches: 4

```

Query Match: 37.13% Indels: 0
DB: 13 Gaps: 0

```

US-10-023-523-44 (1-546) x BQ389733 (1-654)

215	QY	LysLysGlnSerGlnLeuValGlnGlnLysAspHisLeuArgGlyGluHisSerLysAla	234
		::::	
3	DB	AAAAAACAGACCCCGACTTATTCAAGAAAGGATCTACTTCGTATGAGCACAGCAAGGCA	62
235	QY	ValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSer	254
		::::	
63	DB	ATCCTGGCAAGAACGACGCTCGAAGTTTGTGCAGAGAGCTACAGAGACACAACCGTACA	122
255	QY	LeuLysGluGluGlyValGlnArgAlaArgGluGluGluGluLysArgLysGluValThr	274
123	DB	CTPAAAGGAGGAGGAGTACAGAGGCTCGTGAGGAGGAGAGAACGCCAAGAAGTATACC	182
275	QY	SerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArg	294
183	DB	TCCCATTTCCAAAGTGACGCTCAATGATCATCCAGTCCAGATGGAGCAGCACAAATGAACGT	242
295	QY	AsnSerLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGlu	314
243	DB	AATGCTTAACTCGCTCAAGAGAAATGTTGAGCTTGCTGATCGATTAAGAAGCTTATCGAG	302
315	QY	GlnTyrGluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGln	334
303	DB	CAATATGAGCTTAAGAGAGAGACACATTCACAAAGTTTCAACATTAAGGATCTTCAACAG	362
335	QY	GlnLeuValAspAlaLysLeuGlnGlnAlaGlnGlnMetLeuLysGluAlaGluGluArg	354
363	DB	CAACTGTGTGATGCGAGCTTCACACAGCACAGAANTGCTGAAGAGAGTGGAGAGCGT	422
355	QY	HisGlnArgGluLysAspPheLeuLysGluAlaValGlnSerGlnArgMetCysGlu	374
423	DB	CACCAGCGGGAAAAGGAGTTTCTGTTGAAAGAGGAGCAGTCGAATCCCGCGCATGTGTGAA	482
375	QY	LeuMetLysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPhe	394
483	DB	CTAATGAACACGACAGAGACCCACCTCAACACACAGTTACCACTGTATACAGAGAAGTTT	542
395	QY	GluGluPheGlnAenThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGlu	414
543	DB	GAGGAGTTTCAAAACACCCCTGCTCTAAAGCAATGAGGTTTTTCCACCTTTTAAACAAGAA	602
415	QY	MetGluLysMetThrLysLysIleLysLysLeuGluLysGluThrThrMet	431
603	DB	ATGGAAAAGATGACAAAGACATTAAAGAGCTGGAAAAGAGACACACCATG	653

RESULT 44
BU151934
LOCUS
DEFINITION
BU151934 linear mRNA 954 bp
AGENCOURT_8753209 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6394355
5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BU151934
AGENCOURT_8753209 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6394355
5', mRNA sequence.
BU151934
BU151934.1 GI:22665466
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 954)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs@email.nih.gov
Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman,
Ph.D.
cDNA Library Preparation: ResGen, Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAMI3887 row: e column: 12
 High quality sequence stop: 512.

FEATURES	SOURCE
-----------------	---------------

```

/organism="Mus musculus"
/mol_type="mrna"
/db_xref="taxon:10090"
/clone="IMAGE:6394355"
/lab_host="DHIOB (phage-resistant)"

```

ORIGIN

Alignment Scores:	
Pred. No.:	2,38e-55
Score:	1014.50
Percent Similarity:	87.59%
Best Local Similarity:	85.71%
Query Match:	66.43%
DB:	13
Length:	354
Matches:	228
Conservative:	5
Mismatches:	20
Indels:	19
Gaps:	2

US-10-023-523-44 (1-546) X BU151934 (1-954)

291	His	Leu	Glu	Arg	Asn	Ser	Lys	Leu	Arg	Gln	Gln	Leu	Asn	Met	Glu	Leu	Ala	Glu	Arg	Leu	Lys	310	
1	CAT	AAC	GAG	CGA	AACT	CCA	AGCT	CGC	CAG	GAA	TAT	GAG	CTA	CC	CAG	AGG	CT	CAAG	60				
311	Lys	Leu	Ile	Glu	Gln	Tyr	Glu	Leu	Arg	Glu	Glu	His	Ile	Asp	Lys	Val	Phe	Lys	His	Lys	330		
61	AAG	TGT	ATC	GAG	CAAT	ACG	AGC	TCG	TAG	GAG	CA	TAT	CGA	CAA	ATC	TT	CAAA	CATA	AG	120			
331	Asp	Leu	Gln	Gln	Gln	Leu	Val	Asp	Ala	Lys	Leu	Gln	Gln	Ala	Leu	Val	Glu	Met	Leu	Lys	350		
151	GAC	TTC	GAC	GAC	GAC	TAG	TGG	AC	CGC	CA	AGCT	CCG	CAG	CGC	CC	CAG	GAG	ATC	TGT	GA	AGG	180	
351	Ala	Glu	Glu	Arg	His	Gln	Arg	Glu	Lys	Asp	Phe	Leu	Lys	Glu	Ala	Val	Glu	Ser	Gln	370			
181	CGA	GAG	GAG	CGC	CA	CC	ACG	CAG	GAG	AGG	AT	TT	CT	CT	CT	GAA	GAA	CGC	TGG	AGT	CC	CA	240
371	Arg	Met	Cys	Glu	Leu	Met	Lys	Gln	Gln	Glu	Thr	His	Leu	Lys	Gln	Gln	Leu	Ala	Leu	Tyr	390		
241	AGG	ATG	TCC	GAG	CT	GAT	GAT	GAC	GAG	CAG	CA	CC	CT	TAA	GAC	GAC	CT	CGC	CT	G	TAC	300	
391	Thr	Glu	Lys	Phe	Glu	Glu	Phe	Gln	Asn	Thr	Leu	Ser	Lys	Ser	Ser	Glu	Val	Phe	Thr	Thr	410		
301	ACG	GAG	AGT	TTG	GAG	AGT	TCC	GAA	CAC	ACT	TT	TCC	AAA	AGC	AGT	CA	AGT	TGT	T	CA	CCG	360	
411	Phe	Lys	Gln	Glu	Met	Glu	Lys	Met	Thr	Lys	Lys	Ile	Lys	Lys	Leu	Glu	Lys	Glu	Thr	Thr	430		
361	TT	CAA	CAG	GAG	ATG	GAA	NAG	ATC	CAC	AAA	GAG	ATC	CA	GAA	AGT	TG	GAG	AAA	GAAA	CA	CC	420	
431	Met	Tyr	Arg	Ser	Arg	Tyr	Glu	Ser	Ser	Asn	Lys	Ala	Ileu	Leu	Glu	Val	Ala	Glu	Glu	Lys	450		
421	ATG	TAT	CGAT	CCCG	TGG	AAA	GC	CA	CA	AGG	CT	CT	CT	GAG	ATG	CT	CT	GAG	AGG	AAA	480		
451	Thr	Val	Arg	Asp	Lys	Glu	Leu	Glu	Gly	Lys	Ile	Gln	Val	Lys	Ile	Gln	Arg	Leu	Glu	Lys	470		
481	ACG	TC	-	CGG	CAA	-	GAG	CT	G	GAG	GGC	CT	C	AGT	G	AAA	ATC	-	CAG	CGG	CT	G	537
471	Cys	Arg	Ala	Leu	Gln	Thr	Glu	Arg	Asn	Asp	Leu	Asn	Lys	Arg	Val	Gln	Asp	Leu	Ser	Ala	490		
538	TGC	CGAC	CACT	TGC	AGAC	CCG	AGC	GCAT	GACT	CT	CAC	CAG	AGG	GGT	TAC	GAC	CACT	GACT	GC	A	597		
491	Gly	Gly	Gln	Gly	Ser	Leu	Thr	Asp	Ser	Gly	Pro	Glu	Arg	Arg	Pro	Glu	-----	506					
598	GGG	GGC	-----	AT	CACT	GC	AT	TGG	CT	CT	GAG	CGG	AG	-	CCAGAG	GC	CACT	GGCT	647				
507	-----	Gly	Pro	Gly	Val	Ala	Gln	Ala	Pro	Ser	Ser	Pro	Arg	Val	Thr	Glu	520						

Db 648 CAAGGAACAGGGGNNCTGAAGTCTCGGGCTCACCC-AGCAGCTCTCCNAGGC-ACAGAC 705
 QY 521 AlaProCysThrProGluAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGlu 540
 Db 706 GCTCCTTGCTGCTCAGGAGCCCGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAG 765
 QY 541 ProThrSerAlaAlaGala 546
 Db 766 CCCACCCCTGCACCTGCC 783
 RESULT 45
 CB247310 777 bp mRNA linear EST 09-JUL-2003
 LOCUS UI-M-FIO-cdz-g-06-0-UI.r1 NIH_BMAP_F10 Mus musculus cDNA clone
 DEFINITION IMAGE:6836479 5', mRNA sequence.
 ACCESSION CB247310
 VERSION CB247310.1 GI:28368954
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 777)
 NIH-MGC <http://mgc.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)
 Seq primer: pYX-5.
 FEATURES
 source Location/Qualifiers
 1..777
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6836479"
 /tissue_type="whole brain"
 /dev_stage="embryo 12.5dpc"
 /lab_host="PH10B (T1 phage resistant)"
 /clone_lib="NIH BMAP F10"
 /notes="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according
 Bonaudo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with NotI and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the PolyA tail
 is CAGCCACGAC. This library was created for the University
 of Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."

ORIGIN

Alignment Scores:

Pred. No.: 6,99e-55 Length: 777
 Score: 1005.50 Matches: 209
 Percent Similarity: 86.75% Conservative: 7
 Best Local Similarity: 83.94% Mismatches: 30
 Query Match: 36.10% Indels: 4

DBS: 14 Gaps: 3

US-10-023-523-44 (1-546) x CB247310 (1-777)

QY 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerSerPro 20
 Db 30 ATGAAGAACCCAGACAAAGAAAGTGGGCTCTCCAAACACTCCAACTCGAAGGGAGGCCG 89
 QY 21 GlyGlnProGluAlaGlyProGluGluValaGlnGluArgProSerGlnAlaAlaProAla 40
 Db 90 GGGCAACCGGAAGCAGGAGCCCGAGGAGCCCATGGACGACCCAGACAGACAGCTCTCTGGG 149
 QY 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
 Db 150 GCGGAAGCCGGAAGT---TCCACAAGCCAGGCTCTCGGAAGACCGAGGGGCTCGAGCT 206
 QY 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
 Db 207 AAGCAGCTCAGCTTGGGCTCTCTGTGACCTCTCTGAGGAGCTGAGCGGCTGTTGAA 266
 QY 81 AspileLeuSerThrTyrcysValAspAsnAsnGlnGlyProGlyGluAspGlyAla 100
 Db 267 GACATCTCTGAGTACATCTGTGTGGACAACTCAGGAGGCCCGCTGAGGAGGGAGCA 326
 QY 101 GlnGlyGluProAlaGluProGluAspAlaGlnLysSerArgThrTyrrValAlaArgAsn 120
 Db 327 CAGGTGAGCCCTCTGAGCCGGAAGACACGAGGAAGTCCCGAACCTATGACGCCAGGAAT 386
 QY 121 GlyGluProGluPro---ThrProValValAsnGlyGluLysGluProSerLysGlyAsp 139
 Db 387 GGGGAGCTTGAAACCCAGGACATTCAGTCTGTAACCGGAGGAGAGACCTCTAAGGGAGAG 446
 QY 140 ProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAspHisArgArgPro 159
 Db 447 CCTGGAAACAGAGGAGATCCGAGCGAGTGTAGTGGAGACCGAGACCATCGGAGGCCA 506
 QY 160 GlnGluLysLysLysAlaLysGlyLysGluLysGluLysGluLysGluLysGluLys 179
 Db 507 CAGGAG 566
 QY 180 AsnThrLeuSerThrProGluGluLysLeuAlaLeuAlaLeuLysLysLysLysLys 199
 Db 567 AACACGCTGAGTACCCAG 626
 QY 200 LeuGluGluHisArgAsnSerGlnLysGlnMetLysLysLysLysLysLysLysLys 219
 Db 627 CTGGAAGAGCATCGGAACTCGCAAGAGCAGATGAAGTCTCTGCAAGAGAGAGAGAGAG 686
 QY 220 LeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysLysLysLysLysLys 239
 Db 687 CTCGTGAGGAG 743
 QY 240 LysLeuGluSerLeuLysCysArgGluLeu 248
 Db 744 AG-CTTGAGAGTCTGTGCGGAGAGCTG 769
 RESULT 46
 LOCUS CB247310
 DEFINITION UI-M-FIO-cdb-1-19-0-UI.r1 NIH_BMAP_F10 Mus musculus cDNA clone
 ACCESSION CB247310
 VERSION CB247310.1 GI:24546483
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 780)
 NIH-MGC <http://mgc.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: PYX-5.

FEATURES

source

Location/Qualifiers

```

1..780
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6827396"
/tissue_type="whole brain"
/dev_stage="embryo 13.5, 14.5, 15.5, 16.5, 17.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_PYO"
/note="Organ: Brain; Vector: PYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into PYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCAGACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

```

ORIGIN

Alignment Scores:
 Pred. No.: 7,03e-55 Length: 780
 Score: 1005.50 Matches: 209
 Percent Similarity: 86.75% Conservative: 7
 Best Local Similarity: 83.94% Mismatches: 30
 Query Match: 36.10% Indels: 4
 DB: 14 Gaps: 3

US-10-023-523-44 (1-546) x CA328385 (1-780)

```

QY 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerPro 20
DB 37 ATGAAGAACCAAGACAAAGAAATGGGCTTGCACAACTCCCACTCGAAGGCGACCCG 96
QY 21 GlyGlnProGlnAlaGlyProGluGlyValaGlnLglnArgProSerGlnAlaAlaProAla 40
DB 97 GGGCAACGGAGACAGACCGGAGGAGCCATGGACCCAGACAGACAGCTCTCTGG 156
QY 41 ValGlnAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
DB 157 GCGAAGCCGAAGGT---TCCACAAGCCAGGCTCTCTGGGAAGACCGAGGGGCTCGAGCT 213
QY 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
DB 214 AACACACTAGCTCTGGGGCCCTCTGTGACGTCTCTGAGAGAGCTGAGCCGCAAGTTGGA 273
QY 81 AspileLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAla 100
DB 274 GACATCTGAGTACATCTGTGTGGACAAACATCAGGAGGCCCGGCTGAGGAGGAGCA 333
QY 101 GlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrValAlaAsn 120
DB 334 CAGGGTGAGCCCACTGAGCCCGAAGACACCGAAGTCCCGAACCTATGACGACCGAGAT 393

```

```

QY 121 GlyGluProGluPro---ThrProValValAsnGlyGluLysGluProSerLysGlyAsp 139
DB 394 GGGGAGGCTGAACCCAGGCATTCAGTCGTCAACGGCGAGAAGGAGACCTCTTAAGGGAGAG 453
QY 140 ProAsnThrGluGluLeuLeuArgGlnSerAspGluValGlyAspArgAspHisArgArgPro 159
DB 454 CCTGGAACAGAGGAGATCCGAGCGAGTGATGAAGTTGGAGCCGAGACCATCGAGGCCA 513
QY 160 GlnGluLysLysLysAlaLysGlyLysGluLeuGluLeuThrLeuLeuMetGlnThrLeu 179
DB 514 CAGGAGAGAGAGAGAGAGAGGCTTAGGAGAGAGATCACTCTGCTGATGCGACACTG 573
QY 180 AsnThrLeuSerThrProGluGluLysLeuAlaAlaLeuCysLysLysTyrAlaGluLeu 199
DB 574 AACACGCTGAGTACCCAGAGAGAGAGCTGCTGCACTGTCAAGAGATATGCTGAGCTG 633
QY 200 LeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysGlnSerGln 219
DB 634 CTGGAAGAGCATCGGAATCGCAGAGCAGATGAAGCTCTTGAGAGAGAGAGAGCCAG 693
QY 220 LeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSer 239
DB 694 CTGCTGCAGAGAGAGAGCATCTCGAGGGGAAA---CAGCAGGCTGTCTTGGCCCGAAGC 750
QY 240 LysLeuGluSerLeuCysArgGluLeu 248
DB 751 AG-CTTGAGAGTCTGTGCGGAGCTG 776

```

RESULT 47

BI661586

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BI661586 890 bp mRNA linear EST 12-SEP-2001
 60330571P1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:5351403 5',
 mRNA sequence.

BI661586

BI661586.1 GI:15575822

EST

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 890)

NIH-MGC <http://mgi.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov

Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth

Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1AM11893 row: e column: 04

High quality sequence stop: 828.

FEATURES

source

Location/Qualifiers

1..890

/organism="Mus musculus"

/mol_type="mRNA"

/strain="NNRI"

/db_xref="taxon:10090"

/clone="IMAGE:5351403"

/tissue_type="tumor, gross tissue"

/dev_stage="5 months"

/lab_host="DH10B"

/clone_lib="NCI_CGAP_Mam4"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;

Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.

Library constructed by Life Technologies. Investigators

providing samples: Lothar Hennighausen/Priscilla Furth,

NIH Reference for transgenic model: Li et al., Cell Growth

and Differentiation 7, 3-11 (1996)."

ORIGIN

Alignment Scores:

Pred. No.: 4,86e-54 Length: 890
Score: 993.50 Matches: 208
Percent Similarity: 85.71% Conservative: 8
Best Local Similarity: 82.54% Mismatches: 30
Query Match: 35.67% Indels: 6
DB: 12 Gaps: 3

US-10-023-523-44 (1-546) x B1661586 (1-890)

Qy 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerSerPro 20
Db 129 ATGAGAACCAAGACAAAGAGATGGCCCTGCCAACACTCCAACTCGAAGGCGCAGCCG 188
Qy 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
Db 189 GGGCAACGGGAAGCAGGACCGGAGGAGCCATCGACGCCAGCCAGACAGACAGCTCTCTGGG 248
Qy 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
Db 249 CGGGAACCGGAAGT---TCCACAGCCAGGCTCTCTGGGAAGACCGAGGGGGCTCGAGCT 305
Qy 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
Db 306 AAGCAGCTCAGCGTGGGCGCTCTGTGACGTCTCTGAGGAGCTGAGCGGCGAGTTGGAA 365
Qy 81 AspIleLeuSerThrThrCys-ValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAl 100
Db 366 GACATCTCTGATGATCATCTGCTGTGGACAACTTCAGGGAGGCGCCGCTGAGGAGGAGC 425
Qy 100 aGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrThrValAlaArgAs 120
Db 426 ACAGGGTGAGCCCACTGAGCGCGGAGACAGCGAGAGTCCGAACTTCAGCCAGGAA 485
Qy 120 nGlyGluProGluPro---ThrProValAlaAsnGlyGluLysGluProSerLysGlyAs 139
Db 486 TGGGGAGCCTGAACACGAGCATTCAGTCTGTCACCGCGGAGAGGAGAGCTCTAAGCGAGA 545
Qy 139 pProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAspHisArgArgPr 159
Db 546 GCCTGGAAAGAGAGATCCGACAGTGAAGTGGAGGAGGAGACCATCGAGGCGC 605
Qy 159 oGlnGluLysLysLysAlaLysGlyLeuGlyLysGluLeuThrLeuLeuMetGlnThrLe 179
Db 606 ACAGGAGAAGAAAGAACCCAGGCTCTAGGGAAGGAGATCACTCTGCTGATGCAGACACT 665
Qy 179 uAsnThrLeuSerThrProGluGluLysLeuAlaAlaLeuLysLysLysThrAlaGluLe 199
Db 666 GAACACGCTGAGTACCCAGAGAGAGAGCTGGCTGCACTGTGCAAGAAAGTATGCTGAGCT 725
Qy 199 u-LeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGln-LysLysGlnSer 218
Db 726 GGCTGGAAAGAGATCGGAATCTCGCAGAGCAGATGAAGCTCTTCAGGGAAGAGCAGAGC 785
Qy 219 GlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArg 238
Db 786 CAGCTCTGTGAGGAGAAGACCATCTGCGAGGGGGAACACAGCAGGCTGTCTCTGGGCCGA 845
Qy 239 SerLysLeuGluSerLeuCysArgGluLeu 248
Db 846 GCAAGT---GAAACTGTGTCGGGAGCTG 872

RESULT 48

BF796069
LOCUS 602258906F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4342392 5',
DEFINITION mRNA sequence.
ACCESSION BF796069
VERSION BF796069.1 GI:12101123
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 965)

AUTHORS

NIH-MGC http://mgc.mci.nih.gov/.

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Louis Staudt, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: LLAM9957 row: k column: 01

High quality sequence stop: 557.

FEATURES

Location/Qualifiers

1..965

Source

organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4342392"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH MGC 85"

/note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;

Site 2: SalI; Cloned unidirectionally; oligo-dT primed.

Average insert size 1.867 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 7,21e-54 Length: 965
Score: 991.50 Matches: 227
Percent Similarity: 82.94% Conservative: 16
Best Local Similarity: 77.47% Mismatches: 28
Query Match: 35.60% Indels: 25
DB: 10 Gaps: 2

US-10-023-523-44 (1-546) x BF796069 (1-965)

Qy 275 SerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArg 294
Db 1 TCGAACTTCCAGTCCACTGATGACATTGACATTCAGCTGCAGATGGACACACCAATGAGCGC 60
Qy 295 AsnSerLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuLeuGlu 314
Db 61 AACTCCAAAGCTGCGCAAGAGAACATGGAGCTGGCTGAGAGCTCAAGAGCTGATTGAG 120
Qy 315 GlnTyrGluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGln 334
Db 121 CAGTATGAGCTGCGGAGAGGACATATCCAAAGTCTTCAACACAGAGGACCTACACAG 180
Qy 335 GlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluGluArg 354
Db 181 CAGCTGTGTGATGCCAAGCTCCAGCAGGCCAGGAGATGCTAAAGAGGAGGAGAGCGG 240
Qy 355 HisGlnArgGluLys-AspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysG 374
Db 241 CACCAGCGGAGAGGCGCATTTCTCTGAAAGAGCAGTAGAGTCCCGAGAGATGTGTA 300
Qy 374 uLeuMetLysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPh 394
Db 301 GCTGATGAAGCAGCAGAGAGACCCACCTGAGCAACAGCTTCGCCCTATACACAGAGAACT 360
Qy 394 eGluGluPheGlnAsn-ThrLeuSerLysSerSerGluValPheThrThrPhelysGlnG 414
Db 361 TGAGGAGTCCAGAACACACTTTCCAAAGACGAGGATTCACCACTTCAAGCAGG 420
Qy 414 luMetGluLysMetThrLysLys-IleLysLysLeuGlu-LysGluThrThrMetTyrAr 433

Db 421 AGATGGAAGATGACTAAGAGCATCAAGAGCTGACGAAAGAAACACCATGATACCG 480
Qy 433 gSerArgTTPGluSerSerAnLysAlaLeuLeuGluMetAlaGluGluLysThrValar 453
Db 481 GTCCCGGTGGAGAGACGACACAGGCCCTGCTGAGATGCTGAGGG-GAAACATGTCG 539
Qy 453 gAspLysGluLeuGluGluValLysLysLysLysLysLysLysLysLysLysLys 473
Db 540 GGATACAGAACTGGAGGCTGACAGTCAACAATCCACCGCTGAGAGAGCTGTGCGCGC 599
Qy 473 aLeuGlnThrGluArgAsnLysLeuLysArgValGlnAspLysSerAlaGlyGlyG 493
Db 600 CACTGAAGACAGAGCGGATGACCTGAACAGAGGGGTACG-GACCTGAGTCCCGTGGCCA 658
Qy 493 nGlySerLeuThrAspSerGlyProGluArgArgProGlu-GlyProGlyAlaGlnAlap 513
Db 659 GGGACACCTCACTGACGT-GGCCCTGAGAGAGGCGCAGAGGGGCTGTGGGGCCAG--C 714
Qy 513 roSerSerProArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAla 533
Db 715 CCCAGGCGCCCGAGGTACAGAGCGGTGCTACCCAGGCGCACCGGACAGAGAGCTG 774
Qy 533 erGlyGln-----ThrG 537
Db 775 CAGCCCGAGCGGCTACGAGCCTCCGCGAGGCTACACCTGAGCGCGGAGGAGCG 834
Qy 537 lyProGlnGluProThrSerAlaArgala 546
Db 835 GAGCAGCGGCGACCAAGGCCCGGAGGCA 863

RESULT 49
BU054710 686 bp mRNA linear EST 26-AUG-2002
LOCUS UI-M-FD0-bzi-k-16-0-UI.r1 NIH-BMAP_F00 Mus musculus cDNA clone
DEFINITION IMAGE:6404103 5', mRNA sequence.
ACCESSION BU054710
VERSION BU054710.1 GI:22494787
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 686)
REFERENCE NIH-MGC <http://mgi.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
Location/Qualifiers
1..686
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6404103"
/issue_type="whole brain"
/dev_stage="embryo 12.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH-BMAP_F00"
/note="Organ: Brain; Vector: pYX-Asc; Site: 1: EcoR I;
Site 2: Not I; The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with EcoR I adaptor, digested with Not I, and then
cloned directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA
tail, is TCAGAGAGCC. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP): "Gene Discovery in the Developing Mouse Nervous
System", supported by National Institutes of Mental Health
(NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN
Alignment Scores: 5.92e-54 Length: 686
Pred. No.: 990.00 Matches: 197
Score: 88.21% Conservative: 5
Percent Similarity: 86.03% Mismatches: 25
Best Local Similarity: 35.55% Indels: 2
Query Match: 13 Gaps: 2
DB: US-10-023-523-44 (1-546) x BU054710 (1-686)
Qy 4 GlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerSerProGlyGlnPro 23
Db 2 CAAGACAAAAGAAATGGGCTGCCAAACACATCCAACTCGAAGGGGAGCCCGGGGCAACGG 61
Qy 24 GluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAlaValGluAla 43
Db 62 GAAGCAGAGACCGGAGGAGCCCATGGACGACCCAGACAGACGCTCTCTGGGGCGGAAGCC 121
Qy 44 GluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAlaArgThrAla 63
Db 122 GAAGGT---TCCACAGCCAGCGCTCTCTGGAGAGCCGAGGGGGCTCGAGCTAAAGCAGCT 178
Qy 64 GlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGluAspIleu 83
Db 179 CAGCCTGGGGCCCTCTGTGACGTCTCTGAGGAGCTGAGCCCGGAGTATAGAGACATCTCG 238
Qy 84 SerThrTyrCysValAspAsnGlnGlyGlyProGlyGluAspGlyAlaGluGlyGlu 103
Db 239 AGTACATCTGTGTGACACACATCAGGAGGCCCGCTGAGGAGGAGACACAGGTCAG 298
Qy 104 ProAlaGluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsnGlyGluPro 123
Db 299 CCCACTGAGCCGGAAGACACGAGAAAGTCCCGAACCTATGCGAGCCAGGAATGGGAGGCT 358
Qy 124 GluPro---ThrProValValAsnGlyGluLysGluProSerLysGlyAspProAsnThr 142
Db 359 GAACACGAGATTCAGTCTGTCACCGCGAGAGAGAGAGACTCTAAGGGAGAGCGCTGACACA 418
Qy 143 GluGluIleArgGlnSerAspGluValGlyAspArgAspHisArgArgProGlnGluLys 162
Db 419 GAGGAGATCCGAGCGAGTGTGAAGTTGGAGACCGAGACCATCGGAGGCCACAGAGAGAAG 478
Qy 163 LysLysAlaLysGlyLeuGlyLysGluLeuThrLeuLeuMetGlnThrLeuAsnThrLeu 182
Db 479 AAGAAAGCCAGGGTCTAGGGAGGAGATCATCTGTGATGTCAGACACTGAACACGCTG 538
Qy 183 SerThrProGluGluLysLeuAlaAlaLeuLysLysLysLysLysLysLysLysLys 202
Db 539 AGTACCCACAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 598
Qy 203 HisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysLysLysLysLysLysLys 222
Db 599 CATCGGAACCTCGCAGAGCAGATGAAGCTCTCTGCAAGAGAGAGAGAGAGAGAGAGAG 658
Qy 223 GluLysAspHisLeuArgGlyGluHis 231
Db 659 GAGAGAGACCATCTCGGAGGGGAGAACAC 685
RESULT 50

AY405165 1611 bp DNA linear GSS 12-DEC-2003
 LOCUS Homo sapiens HCM2133 gene, VIRUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.
 ACCESSION AY405165
 VERSION AY405165.1 GI:39761139
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1611)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Perriera,S., Wang,G., Zheng,X.H., White,T.J., Smitsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trices
 JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302
 REFERENCE 2 (bases 1 to 1611)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Perriera,S., Wang,G., Zheng,X.H., White,T.J., Smitsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 source
 1..1611
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 gene
 <1..>1611
 /locus_tag="HCM2133"
 ORIGIN
 Alignment Scores: Length: 1611
 Pred. No.: 2.26e-53 Matches: 206
 Score: 988.00
 Percent Similarity: 61.07% Conservative: 56
 Best Local Similarity: 48.02% Mismatches: 107
 Query Match: 35.48% Indels: 60
 DB: 29 Gaps: 2
 US-10-023-523-44 (1-546) x AY405165 (1-1611)
 QY 176 MetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuLysLysLys 195
 Db 1 ATSCAAATCTGAACAAAGTTGCAACACCCGGAAGAAAGTTTGAATTTTATTCAGAGAG 60
 QY 196 TyrAlaGluLeuGluHisArgAsnSerGlnLysGlnMetLysLeuLysLys 215
 Db 61 TATGCTGAATGCTGTGATGACATCTGCTAGCAAAAGAGTTAAAGCTCTCCAAAG 120
 QY 216 LysGlnSerGlnLeuValGlnLysAspHisLeuArgGlyGluHisSerLysAlaVal 235
 Db 121 AAACAGGTACAAATTCAAAAGAAAGAAAGGACCAAGTTACAAGGTCAACACAGAGATATC 180
 QY 236 LeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisArgSerLeu 255
 Db 181 CTCGCTCGAAGCAAAATGAGAGTCTGTGCCGGAGCTCAGAGACACAAAGACTCTG 240
 QY 256 LysGluGluGlyValGlnArgAlaArgGluGluGluLysGluValThrSer 275
 Db 241 AAGGAAGAGCGCTTCAGCGGGCACCTGAGGAGAGAGAGAAAGGAAGAAATCACAAGC 300
 QY 276 HisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisGlnArgAsn 295
 Db 301 CATTTCCAGATACCTTCACGACATCCAGGGCCAGCATCAGCAGCAGAGTGAAGCAAT 360

QY 236 SerLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLysLeuLysGln 315
 Db 361 ATGAAGCTCTCTCAGGAGAACACAGAGCTTCAGAAAAGCTGAAAAGCATCATCATCAG 420
 QY 316 TyrGluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGln 335
 Db 421 TATGAGCTCAGAGAGGAGCATCTGGCAAAATATTAAACACAGAGACTGCACAGAG 480
 QY 336 LeuValAspAlaLysLeuGlnGlnAlaGlnMetLeuLysGluAlaGluLysGlnHis 355
 Db 481 CTGGTGGATGCAAAAGCTTGACAGGCCCAAGAAATGATGAAGGAAGCGGAGGAGCAGC 540
 QY 356 GlnArgGluLysAspPheLeuLysLysLeuAlaValGluSerGlnArgMetCysGluLeu 375
 Db 541 AAACAGAAAAGGAAATATTCTGAACACGACAGAGTGGAAACTTCAGCCGAAAGTG 600
 QY 376 MetLysGlnGlnThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGlu 395
 Db 601 CTGAAGGAGCAAGACAGACAGTCTCGAGGCTCAGCTCTCTCTACTCAGGAAGGTTTGA 660
 QY 396 GluPheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMet 415
 Db 661 GAATTCAGAGCACACTAACTAAAGCAACAGAGTGTTCGACAGTTCACAAACAGCAATG 720
 QY 416 GluLysMetThrLysLysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArg 435
 Db 721 GACAAANNN 780
 QY 436 TrpGluSerSerAsnLysAlaLeuGluMetAlaGluGluLysThrValArgAspLys 455
 Db 781 NNN 840
 QY 456 GluLeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuCysAtgAlaLeuGln 475
 Db 841 CAATATGAGTCTTTGTGATGAAAATCGGAGGCTAGAGAACCTCTGCCGTGCTTTACAA 900
 QY 476 ThrGluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySer 495
 Db 901 GAAGAGAGAAACGAACCTCCACAAAATAATCAGACGCGAGCAAAATATCTGAAAAGGATGAC 960
 QY 496 LeuThrAspSerGlyProGluArgArgProGluGly----- 507
 Db 961 CAAAGTCAGCAAACTCCGATGAAGAGCCAGAGTCAAAACGTCTCTGTGGATCAAGAGATT 1020
 QY 507 ----- 507
 Db 1021 GACGACAGAGGAGTTAATAGTGTCCAAACCCCGTGAATAATCTGGCCACAGCCTTCATG 1080
 QY 507 ----- 507
 Db 1081 ATAATTATCATCCAGAGTCAACCCCGCACCCAGTCCAAAGAAACCCACCCGAAATAGGC 1140
 QY 508 -----ProGlyAlaGlnAlaProSerSerPro----- 516
 Db 1141 AGTTCTCAGGAGAGTGTGACCGCGCTCTCAGAGGAGCCAGAGCAACCCCTCTGATCCCT 1200
 QY 517 ---ArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGln 535
 Db 1201 TCACGGGATTCAGAGAGTCCCTCCCTCCCTTAATCTCCTCAGGCTGAAGCGGAGGAGC 1260
 QY 536 ThrGlyProGluGluProThrSerAla 544
 Db 1261 AGTGATGCTGAACCTCCTCCTCCAGGCC 1287

Search completed: June 8, 2004, 12:35:46
 Job time : 4023.39 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 8, 2004, 06:37:59 ; Search time 972.245 Seconds
(without alignments)
2561.956 Million cell updates/sec

Title: US-10-023-523-44

Perfect score: 2785

Sequence: 1 MNQDKKGAQKQSNPKSSP.....APSTASGQTGQPTSPARA 546

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2995936 seqs, 2280998010 residues

Total number of hits satisfying chosen parameters: 5991872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10023523/runat_08062004_063750_29986/app_query.fasta_1.1422
-DB=Published Applications NA -QMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=50 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=50 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10023523 -CGN 1 1 810 -runat_08062004_063750_29986
-NCPUP=6 -ICPU=3 -NO_MMAP -LARGESUBQ -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:
- 2: /cgn2_6/prodata/2/pubpna/US06_PCT_NEW_PUB.seq:
- 3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:
- 4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:
- 5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:
- 6: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:
- 7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:
- 8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:
- 9: /cgn2_6/prodata/2/pubpna/US09A_PUBCOMB.seq:
- 10: /cgn2_6/prodata/2/pubpna/US09B_PUBCOMB.seq:
- 11: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq:
- 12: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:
- 13: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:
- 14: /cgn2_6/prodata/2/pubpna/US10A_PUBCOMB.seq:
- 15: /cgn2_6/prodata/2/pubpna/US10B_PUBCOMB.seq:
- 16: /cgn2_6/prodata/2/pubpna/US10C_PUBCOMB.seq:
- 17: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq:
- 18: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:
- 19: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

ALIGNMENTS

RESULT 1

US-09-976-740-46
; Sequence 46, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608

1	2785	100.0	1638	9	US-09-976-740-46	Sequence 46, Appl
2	2785	100.0	1638	13	US-10-023-523-46	Sequence 46, Appl
3	2785	100.0	1638	14	US-10-023-523-46	Sequence 46, Appl
4	2785	100.0	1638	14	US-10-023-523-46	Sequence 46, Appl
5	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
6	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
7	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
8	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
9	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
10	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
11	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
12	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
13	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
14	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
15	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
16	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
17	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
18	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
19	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
20	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
21	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
22	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
23	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
24	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
25	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
26	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
27	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
28	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
29	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
30	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
31	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
32	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
33	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
34	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
35	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
36	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
37	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
38	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
39	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
40	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
41	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
42	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
43	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
44	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
45	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
46	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
47	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
48	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
49	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl
50	2785	100.0	1638	15	US-10-023-523-46	Sequence 46, Appl

;; PRIOR FILING DATE: 1997-11-26
;; PRIOR APPLICATION NUMBER: US 60/031,930
;; PRIOR FILING DATE: 1996-11-27
;; PRIOR APPLICATION NUMBER: US 60/048,547
;; PRIOR FILING DATE: 1997-06-03
;; NUMBER OF SEQ ID NOS: 53
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 46
;; LENGTH: 1638
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (1)...(1638)
US-09-976-740-46

Alignment Scores:

Pred. No.: 7,68e-210 Length: 1638
Score: 2785.00 Matches: 546
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-023-523-44 (1-546) x US-09-976-740-46 (1-1638)

QY 1 MetLysAsnGlnAspLysAsnGlyAlaAlaLysGlnSerAsnProLysSerSerPro 20
DB 1 ATGAAGAACCAAGACAAAGAACCGGGCTGCCAACAATCCAAATCCAAAAGACGCCCA 60
QY 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
DB 61 GGACACCGGAAGCAGAGCCCGAGGAGCCAGAGCGGCCAGCCAGCGGGCTCTGCA 120
QY 41 ValGluAlaGluGlyProGlySerSerGlnAlaAlaProArgLysProGluGlyAlaGlnAla 60
DB 121 GTAGAGCAGAGAGTCCCGGAGCAGCAGCGCTCTCGAAGCGGAGGTGCTCAAGCC 180
QY 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
DB 181 AGACGGCTCAGTCTGGGGCCCTTCGTGTGTCTCTGAGAGGCTGAGCCGCCCACTGGAA 240
QY 81 AspIleLeuSerThrTyrCysValAspAsnGlnGlyGlyProGlyGluAspGlyAla 100
DB 241 GACATACCTAGACACATCTGTGTGGACATAAACCCAGGGGGGCCCGCGAGGATGGGCA 300
QY 101 GlnGlyGluProAlaGluProGluAspAlaGlnLysSerArgThrTyrValAlaAlaArgAsn 120
DB 301 CAGGGTGAGCGCGCTGAACCCGAAGATGCAAGAGTCCCGGACCTATGTGGCAAGGAT 360
QY 121 GlyGluProGluProThrProValValAsnGlyGlyGluProSerLysGlyAspPro 140
DB 361 GGGGAGCTGAACCACTCCAGTAGTCAATGGAGAGAAGAACCCCTCCAAAGGGGATCCA 420
QY 141 AsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHisArgAspProGln 160
DB 421 AACACAGAAGAGATCCGGCAGAGTGACAGTCCGAGACCCAGACCATCGAGGCCACAG 480
QY 161 GluLysLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsn 180
DB 481 GAGAAGAAAAGCCAAAGGGTTTGGTAAGGAGATCACGTTGTGTGTCAGACATTGAAT 540
QY 181 ThrLeuSerThrProGluGluLysLeuAlaAlaLeuCysLysLysTyrAlaGluLeuLeu 200
DB 541 ACTCTGAGTACCCAGAGGAGAGCTGGCTGCTGTGCAAGAGATGATGTAACCTGCTG 600
QY 201 GluGlnHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysGlnSerGlnLeu 220
DB 601 GAGAGGACCCGGAATTCACAGAACCAAGTGAAGTCTCTACAGAAAAGCAGAGCCAGCTG 660
QY 221 ValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLys 240
DB 661 GTCAAGAGAGGAGGACCACTCGCGGGTGAGCACAGCAAGCCGCTCTGGCCCCCAGCAAG 720

RESULT 2

US-10-671-242-46
; Sequence 46, Application US/10671242
; Publication No. US2004004049A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.

QY 241 LeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLysGluGlyVal 260
DB 721 CTTGAGAGCCTATCCGCTGAGCTGAGCGGCACAAACCGCTCCCTCAAGGAAGAGGTGTG 780
QY 261 GlnArgAlaArgGluGluGluGluLysArgLysGluValThrSerHisPheGlnValThr 280
DB 781 CAGCGGCCCCGGAGGAGGAGAGCGCAAGGAGGTGACCTCGCACTTCCAGGTGACA 840
QY 281 LeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGln 300
DB 841 CTGAATGACATTGAGCTGAGATGAAACAGACACAAATGAGCGCAACTCAAGCTCGGCAA 900
QY 301 GluAsnMetGluLeuAlaGluArgGluLysLysLeuIleGluGlnTyrGluLeuArgGlu 320
DB 901 GAGAACATGGAGCTGGCTGAGAGGCTCAAGAGCTGATTGAGCAGTATGAGCTCGGGAG 960
QY 321 GluHisIleAspLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLys 340
DB 961 GAGCATATCGCAAAAGTCTTCAAAACAAAGGACCTACACAGCAGCTGGTGGATGCCAAG 1020
QY 341 LeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGlnArgGluLysAsp 360
DB 1021 CTCAGCAGGCCCCAGGAGATGCTAAAGGAGCAGAGCGGCCACCGCGGAGAGGAT 1080
QY 361 PheLeuLeuLysGluAlaValIleLysSerGlnArgMetCysGluLeuMetLysGlnGln 380
DB 1081 TTTCTCTGAAAGAGGAGGAGTAGAGTCCCGAGAGGATGTGTGAGCTGATCAAGCAGCAAG 1140
QY 381 ThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThr 400
DB 1141 ACCCACTTGAAGCAACAGCTTGCCTATACACAGAAAGTTTGAGGAGTTCAGAAACACA 1200
QY 401 LeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLys 420
DB 1201 CTTTCCAAAGAGCGGAGGTATTCAACACATTCAAGCAGGAGATGGAAGAATGACTAAG 1260
QY 421 LysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArgTyrGluSerSerAsn 440
DB 1261 AAGATCAAGAAGCTGGAGAAAACCAACCATGTACCGGTCCCGGTGGGAGAGCAGCAAC 1320
QY 441 LysAlaLeuLeuGluMetAlaGluLysThrValArgAspLysGluLeuGluGlyLeu 460
DB 1321 AAGGCCCTGCTTGAGATGGCTGAGAGAAACAGTCCGGATTAAGAACTGGAGGGCTG 1380
QY 461 GlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAsp 480
DB 1381 CAGGTAAAAATCCAAAGCGCTGGAGAAAGCTGTGCCGGGCACCTGCAGACAGCGCAATGAC 1440
QY 481 LeuAsnLysArgValGlnAspLeuSerAlaGlyGlnGlnGlySerLeuThrAspSerGly 500
DB 1441 CTGAACAAGAGGGGTACAGGACCTGAGTCTGGTGGCCAGGGCTCCCTCAGTACAGTGGC 1500
QY 501 ProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGlu 520
DB 1501 CCTGAGAGGAGGCCAGAGGGGCTTGGGCTCAAGCACCCAGCTCCCGCAGGGTCAAGAA 1560
QY 521 AlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGlu 540
DB 1561 GCGCCTTCTCTACCCAGGAGCAGGAGCAGGAGCATCAGGAGCATCAGGCCACTTGGCCCTCAAGAG 1620
QY 541 ProThrSerAlaArgAla 546
DB 1621 CCCACTCCGCGCAGGGCC 1638


```

; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/671,242
; CURRENT FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 1638
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1638)
US-10-671-242-46

Alignment Scores:
Pred. No.: 7,68e-210 Length: 1638
Score: 2785.00 Matches: 546
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-023-523-44 (1-546) x US-10-671-242-46 (1-1638)

QY 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerPro 20
DB 1 ATGAAGAACCCAGACAAAAGACGGGGCTCCCAACATCCCAATCCAAAAGACGCCA 60

QY 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaProAla 40
DB 61 GGACAAACCGGAAGACAGGACCCGAGGAGCCAGGAGCCGAGGAGCCGAGGAGGCTCTGCA 120

QY 41 ValGluAlaGluGlyProGlySerSerGlnAlaProAlaGlyProGluGlyAlaGlnAla 60
DB 121 GTAGAGCAGAGAGGTCCCGGACGACCCAGGCTCTCGGAAGCCGGAGGCTCTCAAGCC 180

QY 61 ArgThrAlaGlnSerGlyAlaAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
DB 181 AGAAGCGCTCAGTCTGGGGCCCTTCGTGATGCTCTGAGGAGCTGAGCGCCCACTGGAA 240

QY 81 AspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAla 100
DB 241 GACATACTGAGCACATACTGTGTGGACAATAACCGAGGGGGCCCGCGGAGGATGGGCA 300

QY 101 GlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsn 120
DB 301 CAGGGTGACCGGCTGAACCCGAGATGCGAGAGAGTCCCGGACCTATGTGCGACAGGAT 360

QY 121 GlyGluProGluProThrProValValAsnGlyGluLysGluProSerLysGlyAspPro 140
DB 361 GGGGAGCCCTGAACCACTCCAGTAGTCAATGAGAGAGAGAAACCCCTCCAAAGGGGGATCCA 420

QY 141 AsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHisArgArgProGln 160
DB 421 AACACAGAGAGATCCGGCAGAGTACGAGAGTCCGAGAGCCGAGACCATCGAAGGCCACAG 480

QY 161 GluLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsn 180
DB 481 GAGAGAGAAAAAGCAAGGGTTTGGTAAAGGAGATCACGTTCTCTGATGACAGACATTGAAT 540

```

```

QY 181 ThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysTyrAlaGluLeuLeu 200
DB 541 ACTCTGAGTACCCAGAGGAGAAAGCTGGCTGCTGTGTGCAAGAGTATGCTGAATGCTG 600

QY 201 GluGluHisArgAsnSerGlnLysGluMetLysLeuGlnLysGlnSerGluLeu 220
DB 601 GAGGAGCACCGGAATTACAGAACGATGAAGTCTCTACAGAAAAAGCAGAGCAGCTG 660

QY 221 ValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaAspSerLys 240
DB 661 GTGCAAGAGAGAGGACCACTGCGCGGTGAGCACAGCAAGCGCTCTCTGCGCCGAGCAAG 720

QY 241 LeuGlnSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGlyVal 260
DB 721 CTTGAGAGCCTATGCTGCTGAGCTGCGAGCGGCACAAACCGCTCCCTCAAGGAGAGG 780

QY 261 GlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnValThr 280
DB 781 CAGCGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840

QY 281 LeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGln 300
DB 841 CTGAATGACATTCAGCTGCGATGGAACAGACACCAATGAGCGCAACTCCAAGCTCGCAA 900

QY 301 GluAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTyrGluLeuArgGlu 320
DB 901 GAGAACATGAGGCTGGCTGAGGAGCTCAAGAAGCTGATTGAGCAGTATGAGCTCGCGAG 960

QY 321 GluHisIleAspLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLys 340
DB 961 GAGCATATCGACAAAGCTCTCAACACACAGACGACCTACACAGCAGAGCTGGTGGATGCA 1020

QY 341 LeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGlnArgGluLysAsp 360
DB 1021 CTCAGCAGGCGCCAGGAGATGCTAAAGAGGAGCAGAGCGGCGCCAGCGGAGAGGAT 1080

QY 361 PheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlu 380
DB 1081 TTTCTCTGAAAGAGGAGGAGTAGAGTCCAGAGGATGTGTGAGCTGATGAGCAGCAGAG 1140

QY 381 ThrHisLysLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThr 400
DB 1141 ACCCACCTGAGCAACAGCTTGCCCTATACACAGAGAAGTTTGAGGAGTTCACAGAAC 1200

QY 401 LeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLys 420
DB 1201 CTTTCCAAAAGCAGCAGAGTATTCACCATTTCAAGCAGGAGATGGAAGAATGACTAAG 1260

QY 421 LysIleLysLysLeuGluLysGluThrMetTyrArgSerArgTyrGluSerSerAsn 440
DB 1261 AAGATCAAGAAGCTGGAGAAAGAAACCAACCATGTACCGGTCCCGGTGGGAGAGCAGCA 1320

QY 441 LysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGluLeuGluGlyLeu 460
DB 1321 AAGCCCTGCTTGAGTGGCTGAGGAGAAACAGTCCGGGATAAAGAACTGGAGGGCTG 1380

QY 461 GlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAsp 480
DB 1381 CAGGTAAATCCAAACGCTGGAGAACTGTCCGGGCACATGCGAGCAGAGCGCAATGAC 1440

QY 481 LeuAsnLysArgValGlnAspLeuSerAlaGlyGlnGlnLysSerLeuThrAspSerGly 500
DB 1441 CTGAACAGAGGGTACAGGACCTGAGTGTGGTGGCCAGGGCTCCCTCACTGACAGTGGC 1500

QY 501 ProGluArgArgProGluGlyProGluValAlaGlnAlaProSerSerProArgValThrGlu 520
DB 1501 CTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560

QY 521 AlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGlu 540
DB 1561 GCGCTTGTCTACCCAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1620

```


QY 481 LeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSerGly 500
Db 1441 CTGAACAAGAGGTACAGACCTAGTGTGGTGGCCAGGGTCTCCCTACTGACAGTGGC 1500
QY 501 ProGluArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGlu 520
Db 1501 CCTGAGAGAGGCCACAGGGGCTGGGGCTCAAGCACCCAGCTCCCCACAGGTCACAGAA 1560
QY 521 AlaProCysTyProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGlu 540
Db 1561 GCGCCTTGTTACCCAGAGCACCAGCAGCAGAGCATCATCGGCAGACTGGGCTCAAGAG 1620
QY 541 ProThrSerAlaArgAla 546
Db 1621 CCCACCTCCGCCAGGGCC 1638

RESULT 4

US-10-023-523-46
; Sequence 46, Application US/10023523
; Publication No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001

; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 1638
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1638)
US-10-023-523-46

Alignment Scores:
Pred. No.: 7,68e-210 Length: 1638
Score: 2785.00 Matches: 546
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-10-023-523-44 (1-546) x US-10-023-523-46 (1-1638)

QY 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerSerPro 20
Db 1 ATGAAGAACCACAGACAAAGAACGGGGCTGCCAAACAAATCCCAATCCAAAAGCAGCCCA 60
QY 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
Db 61 GGACACCGAGACAGACCCCGAGGGAGCCACAGAGCGGCCACAGCGGGCTCTGCA 120
QY 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
Db 121 GTAGAAGCAGAAGTCCCGGCAGCAGCGCTCCTCGGAAGCCGGAGGCTGCTCAAGCC 180

QY 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
Db 181 AGAACGGCTCAGTCTGGGGCCCTTCGTGATGTCCTGAGGAGCTGAGCGCCCACTGAA 240
QY 81 AspileLeuSerThrTyrcysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAla 100
Db 241 GACATACCTGAGCACATACCTGTGTGCAATAACACAGGGGGGCCCGGGAGGATGGGCA 300
QY 101 GlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrcysValAlaArgAsn 120
Db 301 CAGGGTGAAGCGGGTGAACCCGAGATGCGAGAAAGTCCCGACCTATGTGCGCAAGAA 360
QY 121 GlyGluProGluProThrProValValAsnGlyGlyLysGluProSerLysGlyAspPro 140
Db 361 GGGAGGCTGAAACCACTCCAGTAGTCAATGAGAGAAAGAACCTCCAGGGGGGATCCA 420
QY 141 AsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAspHisArgArgProGln 160
Db 421 AACACAGAAAGAGATCCGGCAGAGTGAAGAGTCCGAGACCCGAGACCATCCGAAGGCCACAG 480
QY 161 GluLysLysLysAlaLysGlyLysGlyLysLeuLeuLeuThrLeuLeuMetGlnThrLeuAsn 180
Db 481 GAGAAGAAAAAAGCCAGGGTTTGGTAAGAGATCACGTTGCTGATGACACATTTGAT 540
QY 181 ThrLeuSerThrProGluGluLysLeuAlaLeuLysLysLysLysLysLysLysLysLys 200
Db 541 ACTCTGAGTACCCAGAGAGGAAAGCTGGCTCTCTGTGCAAGAAGTATGCTGAATGCTG 600
QY 201 GluGluHisArgAsnSerGlnLysGlnMetLysLysLeuLeuGlnLysLysGlnSerGlnLeu 220
Db 601 GAGGAGCACCGGAATTCACAGAAAGCAGATGAAGTCTCTACAGAAAAAGCAGAGCCAGCTG 660
QY 221 ValGlnGluLysAspHisLeuArgGlyGlyLysLysLysLysLysLysLysLysLysLys 240
Db 661 GTGCAAGAGAAAGACCCCTCGCGGTGAGCAGCAGCAAGCGCTGCTGCTGCTGCTGCTG 720
QY 241 LeuGluSerLeuLysArgGluGluGlnGlnHisAsnArgSerLeuLysGluGlyVal 260
Db 721 CTTGAGAGCCTATGCGGTGAGCTGCGGGCAGCAGCAACCGCTCCCTCAAGGAAGAAGTGTG 780
QY 261 GlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnValThr 280
Db 781 CAGCGGGCCCGGAG 840
QY 281 LeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGln 300
Db 841 CTGAATGACATTCAGCTGACATGCAACAGCACAATGAGCGCAACTCCCAAGCTGCGCAA 900
QY 301 GluAsnMetGluLeuAlaGluArgLysLysLysLysLysLysLysLysLysLysLysLys 320
Db 901 GAGAACTGGAGCTGGCTGAGAGAGCTCAAGAGCTGATTGAGCAGTATGAGCTGCGCGAG 960
QY 321 GluHisLysAspLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLys 340
Db 961 GAGCATATCGCAAAAGTCTTCAAAACACAGAGACCTTACAAACAGCAGCTGGTGGATGCCAAG 1020
QY 341 LeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluLysArgHisGlnArgGluLysAsp 360
Db 1021 CTCCAGCAGGGCCAGGAGATGCTTAAAGGAGCAGAGAGCGGCGCACCCAGCGGAGAGGAT 1080
QY 361 PheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlu 380
Db 1081 TTTCTCTCTGAAAGAGGCGAGTAGATCCAGAGAGATGTGTGAGCTGATGAGCAGCAAGAG 1140
QY 381 ThrHisLeuLysGlnGlnLeuAlaLeuTyrcysValPheGluLysPheGluGluPheGlnAsnThr 400
Db 1141 ACCACCTTGAAGCAACAGCTTGGCTTATACACAGAGAAAGTTTGGAGGATTTCCAGAACACA 1200
QY 401 LeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLys 420
Db 1201 CTTTCAAAGCAGCGAGTATTACCACTTCAAGCAGGAGATGGAAGATGACTAAG 1260

Qy 421 LysileLysLysLeuGluLysGluThrThrMetTyrArgSerArgTrrpGluSerSerAsn 440
 Db 1261 AAGATCAAGAGCTGGAGAAAGAAACACCATGTACCGGTCCCGGTGGGAGAGCAGCAAC 1320
 Qy 441 LysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGluLeuGluGlyLeu 460
 Db 1321 AAGGCCCTGCTTGAGATGGCTGGAGGAAACAGTCCGGATAAAGACTGGAGGCCCTG 1380
 Qy 461 GlnValLysileGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAsp 480
 Db 1381 CAGTAAATCCAAACGCTGGAGAACTGTGCGGGCACTGCAGACAGAGCCCAATGAC 1440
 Qy 481 LeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSerGly 500
 Db 1441 CTGAACAAGAGGTACAGGACCTGAGTGTGGTGGCCAGGGCTCCCTCACTGACAGTGGC 1500
 Qy 501 ProGluArgArgProGluGlyProGlyValAlaGlnAlaProSerProArgValThrGlu 520
 Db 1501 CCTGAGAGAGGCCAGAGGGGCTGGGGCTCAAGCACCAGCTCCCCAGGGTCAAGAA 1560
 Qy 521 AlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGlu 540
 Db 1561 GCGCCTTGCTACCCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1620
 Qy 541 ProThrSerAlaArgAla 546
 Db 1621 CCCACCTCCCGCAGGGCC 1638

RESULT 5
 US-10-616-187-46
 ; Sequence 46, Application US/10616187
 ; Publication No. US20040013668A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lees, Ann M.
 ; APPLICANT: Lees, Robert S.
 ; APPLICANT: Law, Simon W.
 ; APPLICANT: Arjona, Anibal A.
 ; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
 ; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
 ; TITLE OF INVENTION: ATHEROSCLEROSIS
 ; FILE REFERENCE: 10797-004001
 ; CURRENT APPLICATION NUMBER: US/10/616,187
 ; CURRENT FILING DATE: 2003-07-09
 ; PRIOR APPLICATION NUMBER: US/09/616,289
 ; PRIOR FILING DATE: 2000-07-14
 ; PRIOR APPLICATION NUMBER: US 09/517,849
 ; PRIOR FILING DATE: 2000-03-02
 ; PRIOR APPLICATION NUMBER: US 08/979,608
 ; PRIOR FILING DATE: 1997-11-26
 ; PRIOR APPLICATION NUMBER: US 60/031,930
 ; PRIOR FILING DATE: 1996-11-27
 ; PRIOR APPLICATION NUMBER: US 60/048,547
 ; PRIOR FILING DATE: 1997-06-03
 ; NUMBER OF SEQ ID NOS: 53
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 46
 ; LENGTH: 1638
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)...(1638)
 US-10-616-187-46

Alignment Scores:

Pred. No.:	7,69e-210	Length:	1638
Score:	2785.00	Matches:	546
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	16	Gaps:	0

US-10-023-523-44 (1-546) x US-10-616-187-46 (1-1638)

Qy 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerSerPro 20
 Db 1 ATGAAGAACCAAGACAAAGAGAGGGGCTGCCAAACAATCCAATCCAAAAGAGAGCCCA 60
 Qy 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
 Db 61 GGACAAACCGAAGCAGGAGCCCGAGGAGCCCGAGGAGCCCGAGGAGCCCGAGGAGCCCG 120
 Qy 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
 Db 121 GTAGAGCAGAGAAGGTCCCGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 180
 Qy 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
 Db 181 AGAACCGGTCTAGTCTGGGGGCTTCTGATGTCCTGAGGAGCTGAGCCGCAACTGGAA 240
 Qy 81 AspileLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAla 100
 Db 241 GACATATCGACACATCTGTGTGACATATTAACACAGGGGGGCCCCGGGAGGATGGGGCA 300
 Qy 101 GlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsn 120
 Db 301 CAGGGTGAGCGGGTGAACCCGAAAGATGCAGAGAAGTCCCGGAGCTATGTGGCAAGGAAT 360
 Qy 121 GlyGluProGluProThrProValValAsnGlyGluLysGluProSerLysGlyAspPro 140
 Db 361 GGGGAGCCTGAACCAACTCCAGTAGTCAATGAGAGAGAGAACCCCTCCAGGGGGATCCA 420
 Qy 141 AsnThrGluGluLeuLeuArgGlnSerAspGluValGlyAspArgAspHisArgArgProGln 160
 Db 421 AACACAGAAGAGATCCGGCAGAGTACAGAGTCCGAGACCGAGACCATCGAAGGCCACAG 480
 Qy 161 GluLysLysLysAlaLysGlyLeuGlyLysGluLeuLeuLeuLeuLeuLeuLeuLeuLeu 180
 Db 481 CAGAGAAAAAAGCAAGGTTTGGTAAAGAGATCACGTTGCTGATGATGATGATGATGATGAT 540
 Qy 181 ThrLeuSerThrProGluGluLysLeuAlaAlaLeuLysLysLysLysLysLysLysLysLys 200
 Db 541 ACTCTGAGTACCCAGAGAGAGAGAGTGGCTCTGTGCAAGAAGTATGCTGAACTGCTG 600
 Qy 201 GluGluHisArgAsnSerGlnMetLysLysLysLysLysLysLysLysLysLysLysLysLys 220
 Db 601 GAGGAGCACCGGAATTCAGAGAAGCAGATGAAGTCTCTACAGAAAAAGCAGAGCAGCTG 660
 Qy 221 ValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLys 240
 Db 661 GTGCAAG 720
 Qy 241 LeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGluGlyVal 260
 Db 721 CTTGAGAGCCTATGCGGTGAGCTGAGCGGACACACCGCTCCCTCAAGGAAGAGAGGTGTG 780
 Qy 261 GlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnValThr 280
 Db 781 CAGCGGCCCCGGGAG 840
 Qy 281 LeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGln 300
 Db 841 CTGAATGACATTCAGTGCAGATGGAACAGACCAATGAGCGCAACTCCAAGTCCGCCAA 900
 Qy 301 GluAsnMetGluLeuAlaGluArgLeuLysLysLysLysLysLysLysLysLysLysLysLys 320
 Db 901 GAGAACATGGAGCTGGCTGAGAGGCTCAAGAGCTGATTGAGCAGTATGAGTCCGCGCAG 960
 Qy 321 GluHisLysAspLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLys 340
 Db 961 GAGCATATCGACAAAGTCTTCAACACAAAGGACCTACACAGCAGAGTGGTGGATGCCAAG 1020
 Qy 341 LeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGlnArgGluLysAsp 360
 Db 1021 CTCCAGCAGGCCCGAGAGATCTAAAGAGGACAGAGCGGCGGACCGGAGAGAGAT 1080

QY 361 PheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlu 380
Db 1081 TTTCTCTGAAAGAGCAGTAGAGTCCAGAGGATGTGTAGCTGATGAGAGCAGAG 1140
QY 381 ThrHisLeuLysGlnGlnLeuAlaLeuThrThrGluLysPheGluPheGlnAsnThr 400
Db 1141 ACCCACTGAAGCAACAGCTTGCCCTATACACAGAGAGTTTCAGAGATTCACAGAACACA 1200
QY 401 LeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLys 420
Db 1201 CTTTCCAAAGAGCAGGAGTATTCACCATTCACAGAGAGATGAAAGATGACTAAG 1260
QY 421 LysIleLysLeuLysGluLysGluThrThrMetTyrArgSerArgTrpGluSerSerAsn 440
Db 1261 AAGATCAAGAAGCTGAGAGAAAGAACCAACCATGTACCGGTCCCGGTGGAGAGCAGCAAC 1320
QY 441 LysAlaLeuLeuGluMetAlaGluLysThrValArgAspLysGluLeuGluLysLeu 460
Db 1321 AAGGCCCTCTTGAGATGCTGAGGAGAAACACAGTCCGGGATAAAGAACTGGAGGCGCTG 1380
QY 461 GlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAsp 480
Db 1381 CAGGTAAATCAACGGCTGGAGAGCTGTCCGGGCACTGCACAGAGCGCAATGAC 1440
QY 481 LeuAsnLysArgValGlnAspLeuSerAlaGlyGlnGlySerLeuThrAspSerGly 500
Db 1441 CTGAACAAGAGGTACAGGACCTGAGTGTGTGTGGCCAGGGCTCCCTCACTCACAGTGGC 1500
QY 501 ProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGlu 520
Db 1501 CCTGAGAGAGGCCAGAGGGCCCTGGGGCTCAGACACCAGCTCCCCAGGGTCAAGAA 1560
QY 521 AlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGlu 540
Db 1561 GCGCCTTGCTACCCAGAGCAGCAGCAGCAGAGCATCAGGCCAGACTGGGCTCAAGAG 1620
QY 541 ProThrSerAlaArgAla 546
Db 1621 CCACCTCCGCGAGGCC 1638

RESULT 6

US-10-276-774-784
; Sequence 784, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276, 774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 784
; LENGTH: 2523
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-276-774-784

Alignment Scores:

Pred. No.: 1,276-209 Length: 2523
Score: 2785.00 Matches: 546
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-023-523-44 (1-546) x US-10-276-774-784 (1-2523)

QY 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerSerPro 20
Db 119 ATGAGAACCAAGACAAAGAAACGGGGCTGCACAAATCCATCCATCCAAAGAGCAGCCCA 178
QY 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
Db 179 GGACAAACCGGAAGCAGGAGCCCGAGGAGCCAGAGCGGCCCCAGCCAGGCGGCTCTCTGCA 238
QY 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
Db 239 GTAGAGCAGAGAGTCCCGGAGCAGCAGGCTCTCGGAAGCCGAGGGGGCTCAAGCC 298
QY 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
Db 299 AGAACGGCTCAGTCTGGGGCCCTTCGTGATGCTCTCTGAGGAGCTGAGCGCCCAACTGGAA 358
QY 81 AspileLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAla 100
Db 359 GACATACCTAGACATACACTGTGTGGACAAATAACAGGGGGGCCCGCGGAGGATGGGGCA 418
QY 101 GlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsn 120
Db 419 CAGGGTGAGCGCGCTGAACCCGAAGATGCAGAGAAGTCCCGGACCTATGTGGCAAGGAAT 478
QY 121 GlyGluProGluProThrProValValAsnGlyLysGluProSerLysGlyAspPro 140
Db 479 GGGGAGCCTGAACCAACTCCAGTAGTCAATGAGAGAGGAACCCCTCCAGGGGGATCCA 538
QY 141 AsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHisArgArgProGln 160
Db 539 AACACAGAGAGATCCCGCAGAGTGCAGAGTCCGAGACCCGAGACCATCCGAGGCCACAG 598
QY 161 GluLysLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsn 180
Db 599 GAGAAGAAAAAGCAAGGGTTTGGGGAAGGAGATCACGTTGCTGATGTCAGACATTTGAAT 658
QY 181 ThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysTyrAlaGluLeuLeu 200
Db 659 ACTCTGAGTACCCAGAGAGAGAGCTGCTCTCTGTGCAAGAGTATGCTAACTGCTGTG 718
QY 201 GluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeu 220
Db 719 GAGGAGCACCCGGAATTCACAGAGCAGATGAAGCTCTCTACAGAAAAAGCAGAGCCAGCTG 778
QY 221 ValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLys 240
Db 779 GTCAAGAGAGAGGACCACTCGCGGTGAGCAGCAGCAAGCGCTCTCGGCCCGCAGCAAG 838
QY 241 LeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGlyVal 260
Db 839 CTTGAGAGCCTATGCGGTGAGCTGCAGCGGCACACCGCTCCCTCAAGGAAGAGGTGTG 898
QY 261 GlnArgAlaArgGluGluGluGluLysArgLysGluValThrSerHisPheGlnValThr 280
Db 899 CAGCGGGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 958
QY 281 LeuAsnAspileGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGln 300
Db 959 CTGAATGACATTCAGCTGAGATGGAACAGCACAATGAGCGCAACTCCAAGCTCGGCCAA 1018
QY 301 GluAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTyrGluLeuArgGlu 320
Db 1019 GAGAACATGGAGCTGGCTGAGAGGCTCAAGAAGCTGATTGAGCAGTATGAGCTCGCGGAG 1078
QY 321 GluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeuValAspAlaLys 340
Db 1079 GAGCATATCGCAAAAGTCTTCAACACAGAGCACTACACAGCAGCTGGTGGATGCCAAG 1138
QY 341 LeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGlnArgGluLysAsp 360
Db 1139 CTCAGCAGGCGCCAGGAGATCTTAAGAGGAGCAGAGAGCGGCACAGCGGAGAGGAT 1198
QY 361 PheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlu 380

1199	DB	TTTCTCCTGAAAGAGCAGTAGAGTCCAGAGATGTGTGAGCTGATGAAGCAGCAAGAG	1258
381	QY	ThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThr	400
1259	DB	ACCCACCTGNAGCAAAGCTTGCCTATACACAGAGAGATTTCAGAGGTTCCAGAACACA	1318
401	QY	LeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLys	420
1319	DB	CTTTCAAAAGCAGCAGCGATTTTCAACCATTTCAAGCAGGAGATGGAAGAAGATGACTAAG	1378
421	QY	LysIleLysIleLysIleuGluLysGluThrThrMetTyrArgSerArgTrpGluSerSerAsn	440
1379	DB	AGATCAAGAGCTGGAGAAAGAAACACCATGTACCGTCCCGTGGGAGAGCAGCAAC	1438
441	QY	LysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGluLeuGluGlyLeu	460
1439	DB	AGGGCCCTGCTTGAGATGGCTGAGGAGAAACACAGTCCGGGATAAAGAACTGTGAGGGCCCTG	1498
461	QY	GlnValLysIleGlnArgLeuGluLysLysCysArgAlaLeuGlnThrGluArgAsnAsp	480
1499	DB	CAGGTAAATAATCAACGGCTGGAGAGCTGTGCGGGCACTGCAGACAGACGCCAATGAC	1558
481	QY	LeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSerGly	500
1559	DB	CTGAACAAGAGGGTACAGGACCTGAGTGTGTGTGGCCAGGGCTCCCTCACTGACATGGC	1618
501	QY	ProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGlu	520
1619	DB	CCTGAGAGGAGGCACAGGGGGCTTGGGGTCTCAAGCACCCAGCTCCCCCGAGGTCA	1678
521	QY	AlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGlu	540
1679	DB	GGCCCTTGTACCCAGAGGACCCAGCACAGAGCATCAGGCCAGACTGGGCCCTCAGAG	1738
541	QY	ProThrSerAlaArgAla	546
1739	DB	CCACCTCCGACAGGGCC	1756

RESULT 7

US-09-962-055-17
Sequence 17, Application US/09962055
Patent No. US2002052033A1
GENERAL INFORMATION:
APPLICANT: Lees, Ann M.
Lees, Robert S.
Law, Simon W.
Arjona, Anibal A.
TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
TREATING ATHEROSCLEROSIS
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Fastseq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/962,055
FILING DATE: 24-Sep-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/979,608
FILING DATE: 26-NOV-1997
APPLICATION NUMBER: US 60/031,930
FILING DATE: 27-NOV-1996
ATTORNEY/AGENT INFORMATION:

QY 177 GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaAlaLeuCysLysLysTyr 196
DB 483 CAGACATTGATATCTTGAGTACCCAGAGAGAGAGCTGGCTGCTCTGTGCAAGAGTAT 542
QY 197 AlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLys 216
DB 543 GCTGAACCTGCTGGAGGAGCACCAGAAATTCAGAGAGCAGATGAAGCTCCTACAGAAAG 602
QY 217 GlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeu 236
DB 603 CAGAGCCAGCTGGTGAAGAGAGACCACTGGCGGTGAGCAGCAGCAAGCGGCTGCTG 662
QY 237 AlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLys 256
DB 663 GCCGCGACAGCTTGAGAGCTATGCGTGAGCTGCAGCGGCACAAACCGCTCCTCAAG 722
QY 257 GluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHis 276
DB 723 GAAGAAGGTGTGCAGCGCGGCCGAGAGAGAGAGAGCGCAAGAGGTGACCTCGCAC 782
QY 277 PheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSer 296
DB 783 TTCCAGGTGACACTGATGACATTCAGCTGAGATGGAACAGCAATGAGCGCACTCC 842
QY 297 LysLeuArgGlnGluAsnMetGluAlaGluArgLeuLysLysLysLeuLysGluGlnTyr 316
DB 843 AAGCTGCGCCCAAGAGAACATGAGCTGGCTGAGAGGCTCAAGAGAGCTGATTGAGCAGTAT 902
QY 317 GluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeu 336
DB 903 GAGCTGCGGAGGAGCATATCGCAAAAGCTTTCAACACAGAGACCTTACACAGCAGCTG 962
QY 337 ValAspAlaLysLeuGlnGlnAlaGlnMetLeuLysGluAlaGluArgHisGln 356
DB 963 GTGGATGCCAAGCTCCAGCAGCCCGAGAGATGCTAAAGAGGCGAGAGAGCGCACAG 1022
QY 357 ArgGluLysAspPheLeuLeuLysGluAlaValGlnSerGlnArgMetCysGluLeuMet 376
DB 1023 CGGAGAGAGGATTTCTCCTGAAGAGCGAGTAGTCCAGAGAGATGTGTGAGCTGATG 1082
QY 377 LysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGlu 396
DB 1083 AAGCAGCAGAGAGACCCACCTGAGCAACAGCTTGCCCTATACACAGAGATTTGAGGAG 1142
QY 397 PheGlnAsnThrLeuSerLysSerGluValPheThrThrPheLysGlnGluMetGlu 416
DB 1143 TTCCAGAACACACTTTCCAAAGCAGCGAGGTATTCACCACTTCAAGCAGAGATGGAA 1202
QY 417 LysMetThrLysLysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArgTrp 436
DB 1203 AAGATGACTAAGAAAGATCAGAGAGCTGGAGAAAGAAACCACTATGTAACCGTCCCGTGG 1262
QY 437 GluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGlu 456
DB 1263 GAGAGCAGCAACAAAGGCCCTGCTTGAGATGGCTGAGGAGAAACAGCTCCGGGATAAGAA 1322
QY 457 LeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThr 476
DB 1323 CTGAGGCGCTGCGAGGTAAATAATCCAAACGCGCTGAGAGAGCTGTGCCGGGCACTCGACACA 1382
QY 477 GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeu 496
DB 1383 GAGCGCAATGACCTGCAACAGAGGGGTACAGGACCTGAGTCTGTGTGGCAGAGGCTCCCTC 1442
QY 497 ThrAspSerGlyProGluArgArgProGluGlyProGluAlaGlnAlaProSerSerPro 516
DB 1443 ACTGACAGTGGCCCTGAGAGAGGCCAGAGGGCCCTGGGGCTCAAGCAGCCAGCTCCCC 1502
QY 517 ArgValThrGluAlaProCysTyrProGlyValaProSerThrGluAlaSerGlyGlnThr 536
DB 1503 AGGTCACAGAGCGCTTCTTACCCAGAGGACCCGAGCAGAGCATCAGGCCAGACT 1562

QY 537 GlyProGlnGluProThrSerAlaArgAla 546
DB 1563 GGGCCTCAAGAGCCCACTCCCGCCAGGGCC 1592

RESULT 10

US-10-023-529-17
; Sequence 17, Application US/10023529
; Publication No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 4697
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(1592)
US-10-023-529-17

Alignment Scores:
Pred. No.: 4,5e-202 Length: 4697
Score: 2693.00 Matches: 529
Percent Similarity: 99.81% Conservative: 0
Best Local Similarity: 99.81% Mismatches: 1
Query Match: 96.70% Indels: 0
DB: 14 Gaps: 0

US-10-023-523-44 (1-546) x US-10-023-529-17 (1-4697)

QY 17 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 36
DB 3 AAAAGCAGCCAGGACACCGGAAGCAGAGCCCGAGGAGCCCGAGCGGCCAGCCAG 62
QY 37 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 56
DB 63 GCGGCTCTGCACTAGAGCAGAGAGTCCCGCAGCAGCCAGCTCTCCGAGAGCCGGAG 122
QY 57 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 76
DB 123 GCGGCTCAAGCCAGAACCGCTCAGTCTGGGGCCCTTCGTGATGTCTCTGAGAGCTGAGC 182
QY 77 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGly 96
DB 183 CCGCAACTGGAAAGACATCTGAGCAGCATACTGTGGACAATAACCGGGGGGGCCCCGC 242
QY 97 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 116
DB 243 GAGGATGGGGCAGCAGGGTGAGCGGCTGAACCCGGAAGATGCAGAGAGTCCCGGACCTAT 302
QY 117 ValAlaArgAsnGlyGluProGluProThrProValValAsnGlyGluLysGluProSer 136
DB 303 GTGGCAAGAAATGGGGAGCGCTGAACCACTCCAGTAGTCTTATGGAGAGAGAACCCCTCC 362

```
137 LysGlyAspProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAspHis 156
Db 363 AAGGGGATCCAAACACAGAGAGATCGGCAGAGTACGAGGTCCGAGACCCAGACCAT 422
Qy 157 ArgArgProGlnGluLysLysLysAlaLysGlyLeuGlyLysGluLeuThrLeuLeuMet 176
Db 423 CGAAGGCCACAGAGGAAAGAAAGCCAAAGGGTTTGGGGAAGGAGATCACTTTGCTGATG 482
Qy 177 GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLysLysLysLysTyr 196
Db 483 CAGACATTGAATATCTTGAGTACCCAGAGAGAGAGCTGGCTGCTGTGTCNAGAAGTAT 542
Qy 197 AlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLys 216
Db 543 GCTGAACCTGCTGGAGGAGCAGCCGAAATTCAGAGAAGCAGATGAAGCTCTTACAGAAAAG 602
Qy 217 GlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeu 236
Db 603 CAGAGCCAGCTGGTGCAGAGAAAGACACCTGCGCGTGAGCACAAAGGCGGCTCTG 662
Qy 237 AlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLys 256
Db 663 GCCCGCAGCAGCTTGAGAGCTATGCGTGAGCTGCGAGCGGCACAAACCGCTCCCTCAAG 722
Qy 257 GluGluGlyValGlnArgAlaArgGluGluLysArgLysGluValThrSerHis 276
Db 723 GAAGAAGGTGTGCAGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 782
Qy 277 PheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSer 296
Db 783 TTCAGGGTGACACTGAATGACATTCAGCTGCGATGTGAACAGCACATGAGCGCATCC 842
Qy 297 LysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuGluGlnTyr 316
Db 843 AAGCTGCGCCCAAGAACATCGAGCTGGCTGAGAGGCTCAAGAAGCTGATTGACAGTAT 902
Qy 317 GluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeu 336
Db 903 GAGCTGCGCGAGGAGCATATCGAAAGTCTTCAACACAAAGGACCTACAAACAGCAGCTG 962
Qy 337 ValAspAlaLysLeuGlnGlnAlaGlnLeuMetLeuLysGluAlaGluArgHisGln 356
Db 963 GTGATGTCAGAGCTCCAGAGCCAGCAGAGTGTAAAGGAGGAGGAGGAGGAGGAGGAGGAG 1022
Qy 357 ArgGluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMet 376
Db 1023 CCGGAGAAGGATTTCTCTGAAAGAGGAGCAGTAGAGTCCCGAGGAGTGTGTGAGCTGATG 1082
Qy 377 LysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGlu 396
Db 1083 AAGCAGCAGAGACCCACCTGAGACACAGCTTGCCCTATACACAGAGAAGTTTGAGGAG 1142
Qy 397 PheGlnAsnThrLeuSerLysSerGluValPheThrThrPheLysGlnGluMetGlu 416
Db 1143 TTCCAGAACACACTTCCAAAGAGCAGCGAGGTATTACCACTTCAAGCAGGAGATGGAA 1202
Qy 417 LysMetThrLysLysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArgTyr 436
Db 1203 AAGATGACTAAGAGATCAAGAGCTGGAGAAAGAACACACATGTACCGGTCCCGGTGG 1262
Qy 437 GluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGlu 456
Db 1263 GAGAGCAGCAACAAGGCCCTGCTTGAATGGCTGAGGAGAGAAACAGTCCGGGATAAAGAA 1322
Qy 457 LeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThr 476
Db 1323 CTGGAGGGCCCTGCAGGTAAAAATCCACGGCTGGAGAAAGCTGTGCGGGCACTCGACACA 1382
Qy 477 GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlnGlnSerLeu 496
Db 1383 GAGCGCAATGACCTGAACAAGAGGGGTACAGGACCTGAGTGTGGTGGCGAGGGCTCCCTC 1442
```

497 ThrAspSerGlyProGluArgProGluArgProGluArgProGlyAlaGlnAlaProSerSerPro 516
1443 ACTGACAGTGGCCCTGAGAGAGGAGCCAGAGGGGCTGGGGCTCAAGCACCCAGCTCCCC 1502
Qy 517 ArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThr 536
1503 AGGGTCAAGAGAGCGCTTGTACCCAGAGACCCGAGCACAGAGCATCAGGCCAGACT 1562
Qy 537 GlyProGlnGluProThrSerAlaArgAla 546
1563 GGGCCTCAAGAGCCACCTCCGCCAGGGCC 1592

RESULT 11
US-10-023-523-17
; Sequence 17, Application US/10023523
; Publication NO. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1998-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 4697
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(1592)
US-10-023-523-17

Alignment Scores:
Pred. No.: 4,5e-202 Length: 4597
Score: 2693.00 Matches: 529
Percent Similarity: 99.81% Conservative: 0
Best Local Similarity: 99.81% Mismatches: 1
Query Match: 96.70% Indels: 0
DB: 14 Gaps: 0

US-10-023-523-44 (1-546) x US-10-023-523-17 (1-4597)

Qy 17 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnAlaProSerGln 36
Db 3 AAAAGCAGCCAGGACCAACCGGAAGCAGGAGCCAGAGGAGCCAGAGCGGCCAGCCAG 62
Qy 37 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 56
Db 63 GCGGCTCTCTGAGTAGAAGCAAGTCCCGGAGCAGGAGCTCTCGAAGCCGGAG 122
Qy 57 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluLeuSer 76
Db 123 GGGGCTCAAGCCAGACCGCTCAGTCTGGGGCCCTTGTGTGTCTCTGAGGAGCTGAGC 182
Qy 77 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnGlnGlyLysPhe 96
Db 183 CGCCAACTGGAAGACATACCTAGCAGCATACTGTGTGGAACAATACCGAGGGGGCCCCG 242

QY 97 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGlnGlySerArgThrTyr 116
Db 243 GAGGATGGGGCACAGGAGTGGAGCGGTGAACCCGAGATGCAGAGAAGTCCCGACCTAT 302
QY 117 ValAlaArgAsnGlyGluProGluProThrProValValAsnGlyGluGluProSer 136
Db 303 GTGGCAAGGAATGGGAGCCTGAACCACTCCAGTAGTCTATGGAGAGGAACCCCTCC 362
QY 137 LysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHis 156
Db 363 AAGGGGATCCAAACACAGAGAAGATCCGGCAGAGTGACGAGTCCGAGACCGAGACCAT 422
QY 157 ArgArgProGlnGlnLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMet 176
Db 423 CGAAGGCCACAGAGAGAAAAGCAAGGGTTTGGGGAAGGAGATCATCGTTGCTGATG 482
QY 177 GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysTyr 196
Db 483 CAGACATTGATATCTCTGAGTACCCAGAGGAGAGCTGCTCTGTGTCAGAAAGTAT 542
QY 197 AlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLys 216
Db 543 GCTGAATGCTGGAGGAGCACCGGAATTCACAGAGCAGATGAAGCTCTCTACAGAAAAG 602
QY 217 GlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeu 236
Db 603 CAGAGCAGCTGTGCAAGAGAAGGACCCTCGCGGTGAGCACAGCAAGCGCGCTCG 662
QY 237 AlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLys 256
Db 663 GCCCGCAGCAAGCTTGAGAGCCTATCCGCTGAGCTCAGCGGCACAAACCGCTCCCTCAG 722
QY 257 GluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHis 276
Db 723 GAAGAAGGTGTGAGCGGGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 782
QY 277 PheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSer 296
Db 783 TTCAGGTGACATGATGATGACATTCAGCTGAGTGCAGATGCAGACACAAATGAGCGCACTCC 842
QY 297 LysLeuArgGlnGlnAsnMetGluLeuAlaGluArgLeuLysLysLeuLeuGlnTyr 316
Db 843 AAGCTCGCCACAGAGAATGAGCTGCTGAGAGCTCAAGAAGCTGATGAGCAGTAT 902
QY 317 GluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeu 336
Db 903 GAGCTCGCGAGGAGCATATCGCAAGTCTTCAACACAGGACCTCAACAGCAGCTG 962
QY 337 ValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGln 356
Db 963 GTGGATGCCAAGCTCCAGCAGGCCAGGAGATGCTAAAGAGGAGCAAGAGCGGCACACAG 1022
QY 357 ArgGluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMet 376
Db 1023 CGGGAAGAAGATTTCTCTGAAAGAGGAGGAGTATGAGTCCAGAGGATGTCGAGCTGATG 1082
QY 377 LysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGlu 396
Db 1083 AAGCAGCAGAGACCCACCTGAAGCACAGCTTGCTTATACACAGAGATTTGAGGAG 1142
QY 397 PheGlnAsnThrLeuSerLysSerGluValPheThrThrPheLysGlnGluMetGlu 416
Db 1143 TTCCAGAACACACTTTTCCAAAAGCAGCAGGATTTCCACCATTTCAAGCAGGAGATGAA 1202
QY 417 LysMetThrLysLysLysLysLeuGluLysGluThrThrMetTyrArgSerArgTrp 436
Db 1203 AGATGACTTAGAAGATCAAGAGCTGGAGAAGAAACCAACCATGTACCGGTCGCGGTGG 1262
QY 437 GluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGlu 456
Db 1263 GAGAGCAGCAACAAGGCCCTGCTTGATGCTGAGGAGAGAAAACAGTCCGCGGATTAAGAA 1322

QY 457 LeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThr 476
Db 1323 CTGGAGGGCTTCGAGCTTAAATCCACGGCTGGAGAAGCTGTGCGGCGCATCGCAGACA 1382
QY 477 GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeu 496
Db 1383 GAGCCCAATGACCTGAACAAGGGGTACAGGACCTGAGTGTGTGGCCAGGGCTCCCTC 1442
QY 497 ThrAspSerGlyProGluArgArgProGluGlyProGlyValAlaGlnAlaProSerSerPro 516
Db 1443 ACTGACAGTGGCCCTGAGAGGAGGCCAGAGGGCTCGGGCTCAAGCACCAGCTCCCCC 1502
QY 517 ArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThr 536
Db 1503 AGGGTCACAGAAGCGCCTTGCTCCAGGAGCAGCGAGCAGAGCATCAGGCGCAGACT 1562
QY 537 GlyProGlnGluProThrSerAlaArgAla 546
Db 1563 GGGCCTCAGAGAGCCACCTCCGCCAGGGCC 1592

RESULT 12

US-10-616-187-17
; Sequence 17, Application US/10616187
; Publication No. US20040013668A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/616,187
; CURRENT FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 4697
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(1592)
US-10-616-187-17

Alignment Scores:

Pred. No.: 4,5e-202 Length: 4697
Score: 2693.00 Matches: 529
Percent Similarity: 99.81% Conservative: 0
Best Local Similarity: 99.81% Mismatches: 1
Query Match: 96.70% Indels: 0
DB: 16 Gaps: 0

US-10-023-523-44 (1-546) x US-10-616-187-17 (1-4697)

QY 17 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 36
Db 3 AAAAGCAGCCCGCAGCAACCCGAGCAGGAGCCCGAGGAGCCCGAGGAGCCCGCAGCAG 62
QY 37 AlaAlaProAlaValGluAlaGluGlyProGlySerGlnAlaProArgLysProGlu 56
Db 63 CGGCTCTCTGAGTAGAAGCAGAGGTCTCCGCGCAGCAGCCAGGCTCTCTCGAAGCCGAG 122

QY 57 GlyAlaGlnAlaThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 76
DB 123 GGGGCTCAAGCCAGAACGGCTAGTCTGGGGCCCTTCGTGTGTCTCTGAGGAGCTGAGC 182
QY 77 ArgGlnLeuGluAspIleLeuSerThrTyrcysValAspAsnGlnGlyGlyProGly 96
DB 183 CGCAACTGGAGACATCTAGACACATCTGTGTGGACATTAACACAGGGGGCCCGGC 242
QY 97 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluysSerArgThrTy 116
DB 243 GAGGATGGGCGACAGGGTGAAGCCGGCTGAACCCGAAAGATGCAGAGAAGTCCCGGACCTAT 302
QY 117 ValAlaArgAsnGlyGluProGluProThrProValValAsnGlyGluysGluProSer 136
DB 303 GTGGCAAGGAATGGGAGCCTGAAACCACTCCAGTAGTCTATGGAGAGAGGACCCCTCC 362
QY 137 LysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHis 156
DB 363 AAGGGGATCCAAACACAGAGAAGATCCGGCAGAGTGACGAGGTGCGAGACCGAGACCAT 422
QY 157 ArgArgProGlnGluLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMet 176
DB 423 CGAAGGCCACAGAGAGAAAGAAAGCCAGGGTTTGGGGAGGAGATCACGTTGCTGATG 482
QY 177 GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaAlaLeuCysLysLysTy 196
DB 483 CAGACATTGATCTCTGAGTATCCCGCAGAGAGAGAGTGGCTGCTGTGCAAGAAGTAT 542
QY 197 AlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLys 216
DB 543 GCTGAACCTGTGGAGGAGCCCGGAATTCACAGAAGCAGATGAAGCTCCTACAGAAAAAG 602
QY 217 GlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeu 236
DB 603 CAGAGCCAGCTGTGCAAGAGAGAGCCACTCGCGGTGAGGACAGACAGGCGCTGCTG 662
QY 237 AlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLys 256
DB 663 GCCCGCAGCAAGCTTGAGAGCTATGCGGTGAGTGCAGCGGCACACCGCTCCCTCAAG 722
QY 257 GluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHis 276
DB 723 GAAGAAGGTGTGACGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 782
QY 277 PheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSer 296
DB 783 TTCAGGTGACACTGATGATGATTCAGTTCAGTGCAGATGGAACAGCACATGAGCGCAACTCC 842
QY 297 LysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLysLysGlnTy 316
DB 843 AAGCTGGCGCAAGAACATGGAGCTGGCTGAGAGGCTCAAGAGCTGATTGAGCAGTAT 902
QY 317 GluLeuArgGluGluHisLysAspLysValPheLysHisLysAspLeuGlnGlnLeu 336
DB 903 GAGTGGCGGAGGAGCATTCACAAAGTCTTCAACACAGAGGACCTACACAGCAGCTG 962
QY 337 ValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGln 356
DB 963 GTGATGCCAAGCTCCAGAGCCCGCAGAGATGTTAAAGGAGGAGGAGGAGGAGGAGGAG 1022
QY 357 ArgGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMet 376
DB 1023 CGGAGAGAGGATTTCTCTGAAAGAGGAGTAGAGTCCAGAGGATGCTGTGAGCTGATG 1082
QY 377 LysGlnGlnGluThrHisLysLysGlnGlnLeuAlaLeuTyThrGluLysPheGluGlu 396
DB 1083 AAGCAGCAAGAGACCCACCTGAAGCAACAGCTTCCCTATACACAGAGAAGTTGAGGAG 1142
QY 397 PheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGlu 416
DB 1143 TTCAGAACACACTTTCAAAAGAGGAGGAGGATTCACCACTTCAGAGGAGGAGGAGGAG 1202

QY 417 LysMetThrLysLysLysLysLysLeuGluLysGluThrThrMetTyArgSerArgTyr 436
DB 1203 AAGATGACTAAGAGATCAAGAAGCTGAGAGAAAGAACACCACCATGTACCGGTCCCGGTGG 1262
QY 437 GluSerSerAsnLysAlaLeuLeuGluMetAlaGluLysThrValArgAspLysGlu 456
DB 1263 CAGAGCAGCAACAAGGCCCTGCTTGTAGATGCTGAGAGAAACAGTCCCGGATAAAGAA 1322
QY 457 LeuGluGlyLeuGlnValLysLysLysLysLeuGluLysLeuCysArgAlaLeuGlnThr 476
DB 1323 CTGAGGGCCCTGAGGTAAATCCACGCTGGAGAGCTGTGCGGGCACTCGACACA 1382
QY 477 GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeu 496
DB 1383 GAGCGCAATGACCTGAACAAGAGGCTACAGACCTGCTGTGTGCGCAGGGCTCCCTC 1442
QY 497 ThrAspSerGlyProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerPro 516
DB 1443 ACTGACAGTGGCCCTGAGAGGAGCCAGAGGGGCTGGGGCTCAAGCACCCAGCTCCCCC 1502
QY 517 ArgValThrGluAlaProCysTyProGlyAlaProSerThrGluAlaSerGlyGlnThr 536
DB 1503 AGGCTCAGAGAAGCGCTTGTCTACCCAGGAGCACCCGAGCACAGAAGCATCAGGCCAGCT 1562
QY 537 GlyProGlnGluProThrSerAlaArgAla 546
DB 1563 GGGCTCAAGAGCCCACTCCGCCAGGGCC 1592

RESULT 13
US-09-962-055-14
Sequence 14, Application US/09962055
Patent No. US2002005203A1
GENERAL INFORMATION:
APPLICANT: Lees, Ann M.
Lees, Robert S.
Law, Simon W.
Arjona, Anibal A.
TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
TREATING ATHEROSCLEROSIS
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/962,055
FILING DATE: 24-Sep-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/979,608
FILING DATE: 26-NOV-1997
APPLICATION NUMBER: US 60/031,930
FILING DATE: 27-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Louis
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 4722 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

```

; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 61...1731
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-962-055-14
Alignment Scores:
Pred. No.: 5,07e-191 Length: 4722
Score: 2552.50 Matches: 508
Percent Similarity: 93.19% Conservative: 12
Best Local Similarity: 91.04% Mismatches: 25
Query Match: 91.65% Indels: 13
DB: 9 Gaps: 3

US-10-023-523-44 (1-546) x US-09-962-055-14 (1-4722)

QY 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerSerPro 20
DB 61 ATGAAGAATCAAGACAAAAGAACGGGGCTGCCAAGACAGCCCAACCCCAAGACAGCCCG 120
QY 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaProAla 40
DB 121 GGACAGCCGGAAGCAGAGAGCGAGGAGCCAGGGGGCGCCCGCCCGCCCGCCCGCC 180
QY 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
DB 181 CGAAGAGCCGAAGGT---GCCAGACCCAGGCTCCCGGAGGCGCGAGGGGGCTCAAGCC 237
QY 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
DB 238 AAAACTGCTCAGCGCTGGGGCGCTCTGTGTGTCTCTGAGGAGCTGAGCGCCGAGTTGGAA 297
QY 81 AspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAla 100
DB 298 GACATACTAGTACACTGTGTGGACACACACAGGGGGCCCGGGTGGAGTGGGGTC 357
QY 101 GlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsn 120
DB 358 CAGGGTGAGCCCGCTGAACCTGAAGATGACAGAGAAGTCTCGCGCTATGTGCAAGGAAT 417
QY 121 GlyGluProGluPro---ThrProValValAsnGlyGluLysGluProSerLysGlyAsp 139
DB 418 GGGAGCCCGAGCGGGGCGCCCGAGTAGTCAATGGCGAGAAGAGACCTCCCAAGCGCAG 477
QY 140 ProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHisArgArgPro 159
DB 478 CCGGGCACGGAAGAGATCCGGACGACGATGAGTTCGGAGACCGAGACCCCGAGGCCA 537
QY 160 GlnGluLysLysLysAlaLysGlyLeuGlyLysGluIleThrIleLeuMetGlnThrLeu 179
DB 538 CAGGAAAAGAAAGAGCCCAAGGGTCTGGGAAAAGGAGATCACCTGCTGTATGACAGACATG 597
QY 180 AsnThrLeuSerThrProGluGluLysLeuAlaAlaLeuCysLysLysTyrAlaGluLeu 199
DB 598 AACACGCTGAGCACCCCAAGAGAGAGCTGGCGGCTCTGTGCAAGAGTATCGGAACTG 657
QY 200 LeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGln 219
DB 658 CTCGAGGAGCAGCGGAATCTCGCAGAGACAGATGAAGTCTGTCAGAGAAGCAGAGCCAG 717
QY 220 LeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSer 239
DB 718 CTGGTGACGAGAAGAACCACTGCTGGTGAGCAGCAGCAGAGCCATCTGCGCCGCGCAGC 777
QY 240 LysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGly 259
DB 778 AAGCTCGAGAGCTGTGCGCGGAGCTGACAGCGCACCAACCGCTCGCTCAAGGAAGAGGT 837
QY 260 ValGlnArgAlaArgGluGluGluGlnLysArgLysGluValThrSerHisPheGlnVal 279
DB 838 GTGCAGCGAGCCCGCAGAGAGAGGAGAGAGCGCAAGCGAGGTGACGTCACTTCACATG 897

```

```

QY 280 ThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArg 299
DB 898 ACGCTCAACGACATTCAGCTGCAGATGAGCAGCAGCAACAGCGCACTCCAAAGCTGCGC 957
QY 300 GlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTyrGlnLeuArg 319
DB 958 CAGGAGAACATGAGCTGCGCGAGCGCTCAAGAGCTGATTGACAGCTACGAGCTGGCA 1017
QY 320 GluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAla 339
DB 1018 GAAGAGCAGCATCGACAAAGTCTTCAACACACAAGGATCTGCAGCAGCAGCTGGTGACCC 1077
QY 340 LysLeuGlnGlnAlaGlnGlnMetLeuLysGluAlaGluArgHisGlnArgGlnLys 359
DB 1078 AAGCTCCAGAGCGCCAGGAGATGCTGAAGGAGCAGAGGAGCGCACACAGCGGAGAG 1137
QY 360 AspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGln 379
DB 1138 GACTTTCTCTGAAGAGCGCGTGGAGTCCCAGAGGATGTGCGAGCTGATGAGCAACAG 1197
QY 380 GluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsn 399
DB 1198 GAGACCCACCTGAAGCAGCAGCTTGCCTATACACAGAGAAGTTTGAAGAGTTCCAGAAC 1257
QY 400 ThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThr 419
DB 1258 ACTCTTTCCAAAGCAGCAGAGGTGTTCAACACATTCAAACAGGAAATGGAAGATGACA 1317
QY 420 LysLysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArgTyrGluSerSer 439
DB 1318 AAGAAGATCAAGAAGCTGGAGAAAGAGACCAACCATGTACCTTCCCGTGGGAGAGCAGC 1377
QY 440 AsnLysAlaLeuLeuGluMetAlaGluLysThrValArgAspLysGluLeuGluGly 459
DB 1378 AACAGGGCCCTGTGATGAGTGGCTGAGAGAGAAACACTCCGGGCAAGAGACTGGAAGGC 1437
QY 460 LeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsn 479
DB 1438 CTGAGAGTGAAATCCAGCGCTCGAGAGAGCTGTGCGGGCAGCTGCAGACAGAGCGCAAT 1497
QY 480 AspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlnGlySerLeuThrAspSer 499
DB 1498 GACTGTAAACAGAGGGTGCAGAGCTGAGTGTGCGGGCCAGGGGCCCGCTCCGACAGC 1557
QY 500 GlyProGluArgArgPro-----GluGlyPro 508
DB 1558 GGTCTGAGCGGAGCGCAGACCCCGCCACCACTCCAGAGAGCAGGGTGTGAGGGCCCC 1617
QY 509 GlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyrProGlyAlaPro 528
DB 1618 GGGGCTCAAGTACCCAACTCTCCAAAGGGCCACAGACGCTTCTGCTGGCAGGTGCACCC 1677
QY 529 SerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAlaArgAla 546
DB 1678 AGCAGAGGATCAGGCCAGACAGGGGCCCGCCAGAGGCGCCACTGCTCCACTGCCC 1731

```

RESULT 14

US-09-976-740-14
; Sequence 14, Application US/0976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14

; PRIOR APPLICATION NUMBER: US 08/979,608
 ; PRIOR FILING DATE: 1997-11-26
 ; PRIOR APPLICATION NUMBER: US 60/031,930
 ; PRIOR FILING DATE: 1996-11-27
 ; PRIOR APPLICATION NUMBER: US 60/048,547
 ; PRIOR FILING DATE: 1997-06-03
 ; NUMBER OF SEQ ID NOS: 53
 ; SOFTWARE: FastSeq for Windows Version 4.0.
 ; SEQ ID NO 14
 ; LENGTH: 4722
 ; TYPE: DNA
 ; ORGANISM: Oryctolagus cuniculus
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (61)...(1731)
 US-09-976-740-14

Alignment Scores: Length: 4722
 Pred. No.: 5,07e-191
 Score: 2552.50
 Percent Similarity: 93.19%
 Best Local Similarity: 91.04%
 Query Match: 91.65%
 DB: 3

US-10-023-523-44 (1-546) x US-09-976-740-14 (1-4722)

QY 1 MetLysAsnGlnAspTyrLysAsnGlyAlaAlaLysGlnSerAsnProLysSerSerPro 20
 DB 61 ATGAAGAATCAAGACAAAGAGCGGGCTGCGCAACAGCCCAACCCCAAGAGCAGCCG 120
 QY 21 GlyClnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
 DB 121 GGACAGCGGGAAGCAGAGCGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
 QY 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
 DB 181 CGAGAACCCGGAAGT---GCCAGCAGCCAGGCTCCCGGGAGCGCGCGCGCGCGCGCGCG 237
 QY 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
 DB 238 AAAAATCTCTCAGCCTCGGGCGCTCTGATGATCTCTGAGAGCTGAGCCCGCGCGCGCGCG 297
 QY 81 AspIleLeuSerThrTyrCysValAspAsnGlnGlyGlyProGlyGluAspGlyAla 100
 DB 298 GACATCTACTACTACTCTGTGTGCAACACACAGCGCGCGCGCGCGCGCGCGCGCGCGCG 357
 QY 101 GlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsn 120
 DB 358 CAGGCTGAGCCCGCTGAACCTGAAGATGCAGAGAAGTCTCGCGCTATGTGGCAAGGAAT 417
 QY 121 GlyGluProGluPro---ThrProValValAsnGlyGlyLysGluProSerLysGlyAsp 139
 DB 418 GGGGAGCGCGAGCG 477
 QY 140 ProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHisArgArgPro 159
 DB 478 CCGGCGCAGGAAGAGATCCGACGAGCGGATGAGCTCGAGACCCGAGACCCAGCGAGGCCA 537
 QY 160 GlnGluLysLysLysAlaLysGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 179
 DB 538 CAGGAAGAAGAGAGCG 597
 QY 180 AsnThrLeuSerThrProGluGluLysLeuAlaLeuLysLysLysLysLysLysLysLys 199
 DB 598 AACACGCTGAGCACCCAGAGAGAGAGTCTGCGCGCTCTGTGCAAGAAGATATGCGGAATG 657
 QY 200 LeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuGlnLysLysGlnSerGln 219
 DB 658 CTCGAGGAGCAGCGGAACTCCGAGAGCAGATGAAGTCTGCGAGAGAGAGAGAGAGAGAG 717
 QY 220 LeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSer 239

DB 718 CTGGTCAGGAGAGGACCACCTGCGTGGCGCAGCAGCAGCAAGGCCATCTCTGGCCGCGCAGC 777
 QY 240 LysLeuGlnSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGluGly 259
 DB 778 AAGCTCAGAGACCTGTGCGCGGAGCTGACGCGGCACACCCGCTCGCTCAAGAGAGAGGT 837
 QY 260 ValGlnArgAlaArgGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 279
 DB 838 GTGAGGAGCGCGCGAG 897
 QY 280 ThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArg 299
 DB 898 ACGCTCAACGACATTCAGCTGCAGATGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 957
 QY 300 GlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTyrGluLeuArg 319
 DB 958 CAGGAGAACATGGAGCTGGCGAGCGGCTCAGAGAGCTGATTGAGCAGTACGAGGTGCGA 1017
 QY 320 GluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAla 339
 DB 1018 GAAGAGCACATCGACAAAGTCTTCAAAACACAAAGATCTGCAGCAGCAGCTGGTGGACGCC 1077
 QY 340 LysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGlnArgGluLys 359
 DB 1078 AAGCTCCAGCAGCGCCAGGAGATGCTGAAGAGGAGAGAGAGAGAGAGAGAGAGAGAG 1137
 QY 360 AspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGln 379
 DB 1138 GACTTCTCTCGAGAGCGCGCTGAGTCCAGAGGATGTCGAGCTGATGAAGCAACAG 1197
 QY 380 GluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsn 399
 DB 1198 GAGACCCACCTGAAGCAGCAGCTTCCCTATACACAGAGAGATTGAGGAGTCCAGAAC 1257
 QY 400 ThrLeuSerLysSerSerGluValPheThrPheLysGlnGluMetGluLysMetThr 419
 DB 1258 ACTCTTTCCTCAAGAGCGCGAGTGTTCACACATTCACACAGGAATGGAAAGATGACA 1317
 QY 420 LysLysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArgTyrGluSerSer 439
 DB 1318 AAGAAGATCAAGAGCTCGAGAAAGAGAGACCACTATCCGCTCCCGTGGGAGAGCAGC 1377
 QY 440 AsnLysAlaLeuLeuGluMetAlaGluLysThrValArgAspLysGluLeuGluGly 459
 DB 1378 AACAAAGCGCGCTTGAGATGCTGAGAGAGAAACACTCCGCGGACAAAGACTTGAAGGC 1437
 QY 460 LeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsn 479
 DB 1438 CTGCAAGTGAATAATCCAGCGCTGGAGAGAGTGTGCCGCGCACTGCAGACAGCGCAAT 1497
 QY 480 AspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSer 499
 DB 1498 GACCTGAACAGAGAGGTGAGACCTGAGTGGCGTGGCGAGGCGCGCGCTCTCCGACAGC 1557
 QY 500 GlyProGluArgArgPro-----GluGlyPro 508
 DB 1558 GATCTCTGAGCGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1617
 QY 509 GlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyrProGluValaPro 528
 DB 1618 GGGGCTCAAGTACCACTCTCAAGGGGCCAGACGCTTCTGCTGCGCGAGGTGACCC 1677
 QY 529 SerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAlaArgAla 546
 DB 1678 AGCAGAGGCGATCAGGCGAGCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1731

RESULT 15

US-10-671-242-14
 ; Sequence 14, Application US/10671242
 ; Publication No. US20040040049A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lees, Ann M.
 ; APPLICANT: Lees, Robert S.

```

; APPLICANT: Law, Simon W.
; APPLICANT: Ariona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/671.242
; CURRENT FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 4722
; TYPE: DNA
; ORGANISM: Oryctolagus cuniculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)...(1731)
US-10-671-242-14

Alignment Scores:
Pred. No.: 5,07e-191 Length: 4722
Score: 2552.50 Matches: 508
Percent Similarity: 93.19% Conservative: 12
Best Local Similarity: 91.04% Mismatches: 25
Query Match: 91.65% Indels: 13
DB: 13 Gaps: 3

US-10-023-523-44 (1-546) x US-10-671-242-14 (1-4722)

Qy 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerPro 20
Db 61 ATGAAGAATCAAGACAAAAGACGGGCTGCCAAACAGCCCAACCCCAAAAGCAGCCG 120
Qy 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGlnArgProSerGlnAlaAlaProAla 40
Db 121 GGACAGCGGAGAGCAGAGCGGAGGAGCCAGGGCGCGCCCGCGCCCGCCCGCC 180
Qy 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
Db 181 CGAGAGCCGAGGT---GCCAGCAGCCAGGCTCCCGGAGCGCGAGGGGGCTCAAGCC 237
Qy 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
Db 238 AAAAAGTCTCAGCCCTGGGGGCTCTGTGTATGCTCTCTGAGGAGCTGAGCCCGCAGTTGGAA 297
Qy 81 AspileuSerThrTyrcysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAla 100
Db 298 GACATCTAGTACTATCTGTGTGACACACACAGAGGGGCGCCCGCGGTGAGATGGGGTC 357
Qy 101 GlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrcysValAlaArgAsn 120
Db 358 CAGGGTGAAGCCCTGAACCTGAAGATGCAGAGAAGTCTCGCGCCTATGTGGCAAGGAAT 417
Qy 121 GlyGluProGluPro---ThrProValValAsnGlnGlyGluLysGluProSerLysGlyAsp 139
Db 418 GGGAGCGCGGAGCGCGGACCCCAAGTAGTCAATGGCAGAGGAGACCTCCAGGGCAGAG 477
Qy 140 ProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAspHisArgArgPro 159
Db 478 CCGGGCAGCGAGAGATCCCGACCGAGCGATGAGTCCGAGACCGAGACCCACCGGAGGCCA 537
Qy 160 GlnGluLysLysAlaLysGlyLeuGlyGluLeuThrLeuLeuMetGlnThrLeu 179

```

```

Db 538 CAGGAAAAGACAGAGCCCAAGGGTCTGGGAAAGGAGATCACGCTGCTGATGACAGACTG 597
Qy 180 AsnThrLeuSerThrProGluGluLysLeuAlaAlaLeuCysLysLysTyrrAlaGluLeu 199
Db 598 AACAGCTGAGCACCACCCAGAGAGAGCTGGCGGCTGTGTGCAAGAAGTATCGGAAGT 657
Qy 200 LeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGln 219
Db 658 CTCGAGGAGCACCAGGAACTCGCAGAAGCAGATGAAGCTGCTGCAGAGAGCAGAGCCAG 717
Qy 220 LeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSer 239
Db 718 CTGGTGCAGAGAGAGCACCCTGGTGGCGAGCAGCAGCAAGGCCATCTCTCCCGCAGC 777
Qy 240 LysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGly 259
Db 778 AAGCTCGAGAGCTGTGCCGGAGCTGCAGCGGCAACACCGCTCGCTCAGAGAAAGGT 837
Qy 260 ValGlnArgAlaArgGluGluLysArgLysGluValThrSerHisPheGlnVal 279
Db 838 GTGCAGGAGCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 897
Qy 280 ThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArg 299
Db 898 AGCTCAGCAGCATTCAGCTGCAGATGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 957
Qy 300 GlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTyrrGluLeuArg 319
Db 958 CAGGAGAACATGAGCTGCGCGAGCGGCTCAAGAAGCTGATTGACGAGTACGAGCTGCGA 1017
Qy 320 GluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAla 339
Db 1018 GAAGAGCAGCATCGCAAAAGTCTTCAACACACAGAGGATCTGCAGCAGCAGCTGTGTGAC 1077
Qy 340 LysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGlnArgGluLys 359
Db 1078 AAGCTCCAGCAGCCCGAGGAGTGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1137
Qy 360 AspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGln 379
Db 1138 GACTTCTCTCAAGAGGCGCGTGGAGTCCAGAGAGGATGTGCGAGCTGATGAAGCAACAG 1197
Qy 380 GluThrHisLeuLysGlnGlnLeuAlaLeuTyrrThrGluLysPheGluGluPheGlnAsn 399
Db 1198 GAGACCCCTCAAGCAGCAGCTTCCCTTACACAGAGAGTGTGAGGAGTTCAGAGAC 1257
Qy 400 ThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThr 419
Db 1258 ACTCTTTCCAAAAGCAGCGAGGTGTTCCACCATTCAAAACAGGAAATGCAAAAGATGACA 1317
Qy 420 LysLysIleLysLysLeuGluLysGluThrThrMetTyrrArgSerArgTrpGluSerSer 439
Db 1318 AAGAAGATCAAGAAGCTGGAGAAAGAGACCCATGTACCGTTCCTCCGTGGGAGAGCAGC 1377
Qy 440 AsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGluLeuGluGly 459
Db 1378 AACAGGCCCTCTTGAGATGGCTGAGGAGAAAACACTCCCGGACAAAGAGCTGGAGGC 1437
Qy 460 LeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsn 479
Db 1438 CTGCAAGGTGAAATCCAGCGGCTGGAGAGCTGTGCCGGGCACTGCAGACAGAGCGCAAT 1497
Qy 480 AspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSer 499
Db 1498 GACCTGAACAGAGGCTGCAGGACCTGTGTCGGTGGCCAGGGCGCCGCTCTCCACAGCAG 1557
Qy 500 GlyProGluArgArgPro-----GluGlyPro 508
Db 1558 GGTCTCTGAGCGGAGGCCAGAGCGCCACCTCCAAAGGAGCAGGGTGTCTCAGGGCGCCC 1617
Qy 509 GlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyrrProGlyAlaPro 528
Db 1618 GGGGCTCAAGTACCAACTCTCAGAGGGCCACAGACGCTTCTGCTGCGCAGGTGCACCC 1677

```


QY 529 SerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAlaArgAla 546
 DB 1678 AGCACAGAGCATCAGCCACAGAGGGCCCCCAGGAGCCACCACTGCTGCC 1731
 RESULT 16
 US-10-023-529-14
 ; Sequence 14, Application US/10023529
 ; Publication No. US20020129388A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lees, Ann M.
 ; APPLICANT: Lees, Robert S.
 ; APPLICANT: Law, Simon W.
 ; APPLICANT: Arjona, Anibal A.
 ; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
 ; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
 ; TITLE OF INVENTION: ATHEROSCLEROSIS
 ; FILE REFERENCE: 10797-004001
 ; CURRENT APPLICATION NUMBER: US/10/023,529
 ; CURRENT FILING DATE: 2001-12-17
 ; PRIOR APPLICATION NUMBER: 09/616,289
 ; PRIOR FILING DATE: 2000-07-14
 ; PRIOR APPLICATION NUMBER: US 09/517,849
 ; PRIOR FILING DATE: 2000-03-02
 ; PRIOR APPLICATION NUMBER: US 08/979,608
 ; PRIOR FILING DATE: 1997-11-26
 ; PRIOR APPLICATION NUMBER: US 60/031,930
 ; PRIOR FILING DATE: 1996-11-27
 ; PRIOR APPLICATION NUMBER: US 60/048,547
 ; PRIOR FILING DATE: 1997-06-03
 ; NUMBER OF SEQ ID NOS: 53
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 14
 ; LENGTH: 4722
 ; TYPE: DNA
 ; ORGANISM: Oryctolagus cuniculus
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (61)...(1731)
 US-10-023-529-14
 Alignment Scores:
 Pred. No.: 5,07e-191 Length: 4722
 Score: 2552.50 Matches: 508
 Percent Similarity: 93.19% Conservative: 12
 Best Local Similarity: 91.04% Mismatches: 25
 Query Match: 91.65% Indels: 13
 DB: 14 Gaps: 3
 US-10-023-523-44 (1-546) x US-10-023-529-14 (1-4722)
 QY 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerSerPro 20
 DB 61 ATGAAGATATCAAGACAAAGAAAGACGGGGCTGCCAAACAGCCCAACCCCAAGAGCCCG 120
 QY 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
 DB 121 GGACAGCCGGAAGCAGGAGCGAGCGAGCCAGCGGGCGCGCGCGCGCGCGCGCGCGCC 180
 QY 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
 DB 181 CGAGAAGCCGAAGT---GCCAGACAGCAGGCTCCCGGAGCGCGAGGGGGCTCAAGCC 237
 QY 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
 DB 238 AAAACTGCTCAGCTGGGGCGCTCTGTGTCTCTCTGAGGAGCTGACCCGCCAGTTGAA 297
 QY 81 AspLeuLeuSerThrTyrCysValAspAsnAsnGlnGlyProGlyGluAspGlyAla 100
 DB 298 GACATACCTCAGTACATCTGTGTGACAAACAGCGGGCGCGCGGGTGGAGTGGGGTC 357
 QY 101 GlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsn 120

DB 358 CAGGGTGAAGCCCCCTGAACCTGAAGATCGAGAGAGTCTCCGGCTATGTGGCAGGAAT 417
 QY 121 GlyGluProGluPro---ThrProValValAsnGlyGluLysGluProSerLysGlyAsp 139
 DB 418 GGGGAGCCGGAGCGGGCCACCCCTAGTCTCAATGGCGAAGAGGAGACCTCCAAGGCAGAG 477
 QY 140 ProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAspHisArgArgPro 159
 DB 478 CCGGGCACGGAGAGATCCGGACGAGCGATGAGTTCGGAGACCCGAGACCCCGGAGGCCA 537
 QY 160 GlnGluLysLysLysAlaLysGlyLeuGlyLysGluLeuThrLeuLeuMetGlnThrLeu 179
 DB 538 CAGGAAAAGAGAGAGCCAGGGTCTGGAAAAGAGATCACGCTGCTGATCGACACACTG 597
 QY 180 AsnThrLeuSerThrProGluGluLysLeuAlaLeuLysLysLysLysLysLysLysLys 199
 DB 598 AACACGCTGAGCACCCACAGAGAGAGAGCTGGCGGCTCTGTGCAAGAGATATGCGGAAGT 657
 QY 200 LeuGluGluHisArgAsnSerGlnLysGlnMetLysLysLeuGlnLysLysGlnSerGln 219
 DB 658 CTCGAGGAGACCCGGAATCCGAGAGCAGATGAAGTCTGCGAAGAGACGAGCCAG 717
 QY 220 LeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSer 239
 DB 718 CTGGTGCAGGAGAGAGACCCACCTGCTGGCGAGCAGACAGCAAGGCCATCTGGCCGCGAGC 777
 QY 240 LysLeuGluSerLeuLysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGluGly 259
 DB 778 AAGCTCGAGAGCGCTGTGCGGGAGCTGCGGGCGACCAACCCGCTCGCTCAAGAGAAAGGT 837
 QY 260 ValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnVal 279
 DB 838 GTGCGAG 897
 QY 280 ThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArg 299
 DB 898 ACGCTCAACGACATTCAGCTGCGAGATGGAGCAGCACAACAGCGCGCACTCCAAGCTGCGC 957
 QY 300 GlnGluAsnMetGluLeuAlaGluArgLeuLysLysLysLysLysLysLysLysLysLys 319
 DB 958 CAGGAGACATGGAGCTGGCGCGGCTCAAGAGCTGATTGAGCAGTACGAGCTGGGA 1017
 QY 320 GluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAla 339
 DB 1018 GAAGAGCACATCGACAAAGTCTTCAACACAAAGATCTGCGAGCAGCAGCTGGTGACGCC 1077
 QY 340 LysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGlnArgGluLys 359
 DB 1078 AAGCTCCAGCAGCGCCAGGAGATCTGAAGAGGCGAGAGAGCGGACCCAGCGGGAGAG 1137
 QY 360 AspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGln 379
 DB 1138 GACTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1197
 QY 380 GluThrHisLeuLysGlnGlnLeuAlaLeuThrGluLysPheGluGluPheGlnAsn 399
 DB 1198 GAGACCCACCTCAAGCAGCAGCTTCCCTATACACAGAGAGATTGAGAGATTCCAGAAC 1257
 QY 400 ThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThr 419
 DB 1258 ACTCTTTCAAAAGCAGCGAGGTGTTCAACCATTTCAACAGGAAATGGAAAAGATGACA 1317
 QY 420 LysLysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArgTyrGluSerSer 439
 DB 1318 AAGAAGATCAAGAGCTGGAGAAAGAGACCCACCATGTACCGTTCCTCGGTGGAGAGCAGC 1377
 QY 440 AsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGluLeuGluGly 459
 DB 1378 AACAGGCCCTCTTGTAGATGGCTGAGGAGAGAAACACTCCGGGACAAAGAGCTGGAAGGC 1437
 QY 460 LeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsn 479
 DB 1438 CTGCAAGGTGAAATCCAGCGGCTGGAGAGAGCTGTGCGGGCACTGCAGACAGAGCGCAAT 1497


```

Qy 480 AspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSer 499
Db 1498 GACCTGAACAAGAGGGTGCAGGACCTGAGTCCCGGTGCCAGGCGCCCGTCTCCGACAGC 1557
Qy 500 GlyProGluArgArgPro-----GluGlyPro 508
Db 1558 GGTCTGAGCGAGGCCAGAGCCGCCACCACCTCCNAGGAGGAGGGTGTCCGAGGCGCCC 1617
Qy 509 GlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyrProGlyAlaPro 528
Db 1618 GGGGCTCAAGTACCAACTCTCCAAGGGCCACAGACGCTTCTGCTGCGGAGGTGACACC 1677
Qy 529 SerThrGluAlaSerGlyGlnThrGlyProGluGlnProThrSerAlaArgAla 546
Db 1678 AGCACAGGAGCATCAGGCCAGACAGAGGGCCCGCCAGAGCCACCACTGCCACTGCC 1731

RESULT 17
US-10-023-523-14
; Sequence 14, Application US/10023523
; Publication No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Atjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 4722
; TYPE: DNA
; ORGANISM: Oryctolagus cuniculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)...(1731)
US-10-023-523-14

Alignment Scores:
Pred. No.: 5,07e-191 Length: 4722
Score: 2552.50 Matches: 508
Percent Similarity: 93.19% Conservative: 12
Best Local Similarity: 91.04% Mismatches: 25
Query Match: 91.65% Indels: 13
DB: 14 Gaps: 3

US-10-023-523-44 (1-546) x US-10-023-523-14 (1-4722)
Qy 1 MethLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerPro 20
Db 61 ATGAAGAATCAAGACAAAGAACCGGGCTGCCAAACAGCCCAACCCCAAGAGGAGCCCG 120
Qy 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
Db 121 GGACAGCGGAGGAGGAGCGGAGCGCCAGGGCGCGCCCGCGCGCCCGCC 180
Qy 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60

```

```

Db 181 CGAGAAGCGGAGGT---GCCAGACGACGAGGCTCCGGAGGCGCGGAGGGGCTCAAGCC 237
Qy 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
Db 238 AAAACTGTCTCAGCTGGGCGCTCTGTGTATGTCTCTGAGGAGCTGAGCGCCAGTTGGAA 297
Qy 81 AspIleLeuSerThrTyrCysValAspAsnGlnGlyGlyProGlyGluAspGlyAla 100
Db 298 GACATACTCAGTACATACTGTGTGACAAACACAGGGGGGCCCGGGTGAGATGGGGTC 357
Qy 101 GlnGlyGluProAlaGluProGluAlaGluLysSerArgThrTyrValAlaArgAsn 120
Db 358 CAGGCTGAGCCCTGAACTGAAGATGCAGAGAAGTCTCGCGCTATGTGTCAAGGAAT 417
Qy 121 GlyGluProGluPro---ThrProValValAsnGlyGluLysGluProSerLysGlyAsp 139
Db 418 GGGGAGCGGAGCGCGGACCCCTAGTGTCAATGCGGAGAGGAGACCTCCAAAGCGACAG 477
Qy 140 ProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHisArgArgPro 159
Db 478 CCGGGCAGCGAAGAGATCCGAGCAGCGATGAGGTCCGAGACCGAGACACCCGAGGCCA 537
Qy 160 GlnGluLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeu 179
Db 538 CAGGAAAAGAGAGGCGCAAGGCTCTGGGAAAAGGAGATCACGCTCTGTATGCAGACACTG 597
Qy 180 AsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysTyrAlaGluLeu 199
Db 598 AACACGCTGAGCACCAGAGGAGGAGGCTGGCGGCTCTGTGCAAGAGTATCGGAACTG 657
Qy 200 LeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGln 219
Db 658 CTCGAGGAGCAGCGGAACCTCGCAGAAAGCAGATGAAGTCTGTGCAAGAGCAGAGCCAG 717
Qy 220 LeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSer 239
Db 718 CTGGTGCAGAGAGAGGAGCACCTCGTGCAGACAGACAGAGGCCATCTCTGCCCGCCAGC 777
Qy 240 LysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGly 259
Db 778 NAGCTCGAGAGCTGTGCGGGAGCTGCAGCGGCACACCGCTCTCAAGGAAGAGT 837
Qy 260 ValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnVal 279
Db 838 GTGCGAGCGAGCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 897
Qy 280 ThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArg 299
Db 898 AGCTCAACGACATTCAGCTGCAGATGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 957
Qy 300 GlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTyrGluLeuArg 319
Db 958 CAGGAGAACATGGAGCTGGCGGCTCAAGAGGCTGATTGAGCAGTACGAGCTGCGCA 1017
Qy 320 GluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeuValAspAla 339
Db 1018 GAAGAGCAGCATCGACAAAGTCTTCAAAACACAAAGGATCTGCAGCAGCAGCTGTGTGAGCC 1077
Qy 340 LysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGlnArgGluLys 359
Db 1078 NAGCTCCAGCAGCGCCAGGAGATGCTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1137
Qy 360 AspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGln 379
Db 1138 GACTTTCTCTGAAAGAGGCGCGTGGAGTCCGAGAGGATGTGCGAGCTGATGAAGCAACAG 1197
Qy 380 GluThrHisLysLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsn 399
Db 1198 GAGACCCACTGAAGCAGCAGCTTCCCTTATACACAGAGAGAGTTCGAGGAGTTCAGAAC 1257
Qy 400 ThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThr 419
Db 1258 ACTCTTTCAAAAGAGGAGGAGTGTTCACCACTTCAAAACAGGAAATGAAAGATGACA 1317

```

QY 420 LysLysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArgTrrGluSerSer 439
 DB 1318 AAGAAGATCAAGAAAGCTCGAGAAAGAGACCAACCATGTACCGTTCCCGGTGGAGAGCAGC 1377
 QY 440 AsnLysAlaLeuLeuGluMetAlaGluLysThrValArgAspLysGluLeuGluGly 459
 DB 1378 AACAGGCGCCCTGTGAGATGGCTGAGAGAAACACTCCGGGACAAAGAGCTGGAGGC 1437
 QY 460 LeuGlnValLysIleGlnArgLeuGluLysLeuGluLysCysArgAlaLeuGlnThrGluArgAsn 479
 DB 1438 CTGCAGGTGAATAATCCAGGCGCTCGAGAGCTGTGTCCCGGCACCTGCACAGACAGCGCAAT 1497
 QY 480 AspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSer 499
 DB 1498 GACCTGAACAAAGAGGGTGCAGGACCTGAGTCCCGGTGGCCAGGCGCCCTCCGACAGC 1557
 QY 500 GlyProGluArgArgPro-----GluGlyPro 508
 DB 1558 GGTCTGAGCGAGCGCAGACCCCGCCACACCTCCAGGAGCAGGGTGTGAGGGCCCC 1617
 QY 509 GlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyrProGlyAlaPro 528
 DB 1618 GGGGCTCAAGTACCACTCTCCAGGCGCCACAGACGCTTCCTGCTCGGAGGTGCACCC 1677
 QY 529 SerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAlaArgAla 546
 DB 1678 AGCAGAGGCGATCAGGCCAGACAGGGGCCCGCAGGAGCCACCACTGCCACTGCC 1731
 RESULT 18
 US-10-616-187-14
 ; Sequence 14, Application US/10616187
 ; Publication No. US20040013668A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lees, Ann M.
 ; APPLICANT: Lees, Robert S.
 ; APPLICANT: Law, Simon W.
 ; APPLICANT: Arjona, Anibal A.
 ; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
 ; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
 ; TITLE OF INVENTION: ATHEROSCLEROSIS
 ; FILE REFERENCE: 10797-004001
 ; CURRENT APPLICATION NUMBER: US/10/616,187
 ; CURRENT FILING DATE: 2003-07-09
 ; PRIOR APPLICATION NUMBER: US/09/616,289
 ; PRIOR FILING DATE: 2000-07-14
 ; PRIOR APPLICATION NUMBER: US 09/517,849
 ; PRIOR FILING DATE: 2000-03-02
 ; PRIOR APPLICATION NUMBER: US 08/979,608
 ; PRIOR FILING DATE: 1997-11-26
 ; PRIOR APPLICATION NUMBER: US 60/031,930
 ; PRIOR FILING DATE: 1996-11-27
 ; PRIOR APPLICATION NUMBER: US 60/048,547
 ; PRIOR FILING DATE: 1997-06-03
 ; NUMBER OF SEQ ID NOS: 53
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 14
 ; LENGTH: 4722
 ; TYPE: DNA
 ; ORGANISM: Oryctolagus cuniculus
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (61)...(1731)
 US-10-616-187-14

Alignment Scores:

Pred. No.:	5,07e-191	Length:	4722
Scores:	2552.50	Matches:	508
Percent Similarity:	93.19%	Conservative:	12
Best Local Similarity:	91.04%	Mismatches:	25
Query Match:	91.65%	Indels:	13
DB:	16	Gaps:	3

US-10-023-523-44 (1-546) x US-10-616-187-14 (1-4722)
 QY 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerSerPro 20
 DB 61 ATGAAGAATCAAGACAAAAGAACCGGGCTGCCAAACAGACCCCAACCCCAAGAGAGCCG 120
 QY 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
 DB 121 GGACAGCGGAG 180
 QY 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
 DB 181 CCAGAAAGCCGAAGGT---GCCAGGAGCCAGAGCTCCCGGGAGGCGCGAGAGGGGGCTCAAGCC 237
 QY 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
 DB 238 AAAACTGCTCAGCTGGGGCGCTCTGTGATGTCCTCTGAGGAGCTGAGCCCGCAGTTGGAA 297
 QY 81 AspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyProGlyGluAspGlyAla 100
 DB 298 GACATACTCAGTACATACATCTGTGGACAAACAGAGGGGGCGCGGGGTAGGATGGGGTC 357
 QY 101 GlnGlyGluProAlaGluProGluAspAlaGlnLysSerArgThrTyrValaArgAsn 120
 DB 358 CAGGGTGAAGCCCCCTCAACCTGAAGATCGAGAGAGTCTCGCGCTATGTGGCAAGGAAT 417
 QY 121 GlyGluProGluPro---ThrProValValAsnGlyGluLysGluProSerLysGlyAsp 139
 DB 418 GGGGAGCGGAGCGGGCGGCACCCAGTAGTCAATGGCGAGAGGAGACCTCCAGGCGAG 477
 QY 140 ProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHisArgArgPro 159
 DB 478 CCGGGCAGCGAAGAGATCCGACAGAGCATGAGTTCGAGAGCCGAGACACCCAGGAGGCCA 537
 QY 160 GlnGluLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeu 179
 DB 538 CAGGAAAGAGAGAGCCAGGGTCTGGAAAGGAGATCAAGCTGCTGATCGACACTG 597
 QY 180 AsnThrLeuSerThrProGluLysLeuAlaAlaLeuCysLysLysLysLysLysLys 199
 DB 598 AACACGCTGAGCACCCAG 657
 QY 200 LeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuGlnLysLysLysSerGln 219
 DB 658 CTCGAGGAGCACCGGAATCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 717
 QY 220 LeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSer 239
 DB 718 CTGGTGCAGGAG 777
 QY 240 LysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGly 259
 DB 778 AAGCTCGAGAGCTGTGCGGAGCTGAGCGGCAACACCGCTCGCTCAAGAGAGAGAGT 837
 QY 260 ValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnVal 279
 DB 838 GTGCGAGCGAGCGCGAG 897
 QY 280 ThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArg 299
 DB 898 AGGCTCAACGACATTCAGTGCAGATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 957
 QY 300 GlnGluAsnMetGluLeuAlaGluArgLeuLysLysLysLysLysLysLysLysLys 319
 DB 958 CAGGAGAACATGGAGCTGGCGAGCGGCTCAAGAGAGCTGATTGAGCAGTACGAGCTGGGA 1017
 QY 320 GluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAla 339
 DB 1018 GAAGAGACATCGACAAAGTCTTCAACACAGAGGATCTGCAGAGCAGCAGCTGGTGACGCC 1077
 QY 340 LysLeuGlnAlaGlnClnMetLeuLysGluAlaGluArgHisGlnArgGluLys 359
 DB 1078 AAGCTCCAGCAGCGCCAGAGATGCTGAAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1137

```
QY 360 AspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGln 379
Db 1138 GACITTCCTCCTAGGAGGCGCTGGAGTCCAGAGAGATGCGAGCTGATGAAGCAACAG 1197
QY 380 GluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsn 399
Db 1198 GAGACCCACCTGAAAGCAGCAGCTTGCCTATACACAGAGAAGTTTGAGGAGTTCCAGAAC 1257
QY 400 ThrLeuSerLysSerGluValPheThrThrPheLysGlnGluMetGluLysMetThr 419
Db 1258 ACTCTTTCCAAACAGCAGGAGGTGTTCCACACATTCACAGGAAATGGAAGAGATGACA 1317
QY 420 LysLysLeuLysLysLeuGluLysGluThrThrMetTyrArgSerArgTTPGluSerSer 439
Db 1318 AAGAAGATCAAGAAGCTGGAGAAAGACACACCATGTACCGTTCCCGTGGAGAGCAGC 1377
QY 440 AsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGluLeuGluGly 459
Db 1378 AACAGGCCCTGCTTGAGATGGCTGAGGAGAAACACTCCGGGACAAAGAGCTGGAAGGC 1437
QY 460 LeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsn 479
Db 1438 CTGCAGGTGAAATCCAGCGCTGGAGAGCTGTGCCGGGCACTGCAGACAGAGCGCAAT 1497
QY 480 AspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlnGlySerLeuThrAspSer 499
Db 1498 GACCTGAACAGAGGGTGCAGGACCTGAGTGCCTGCTGGCGGCGCCCGTCTCCGACAGC 1557
QY 500 GlyProGluArgArgPro-----GluGlyPro 508
Db 1558 GGTCTGAGCGGAGGCAGAGCCGCCACCACTCCAGAGAGCGGTGTGAGAGGCCCC 1617
QY 509 GlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyrProGlyAlaPro 528
Db 1618 GGGGCTCAAGTACCAACTCTCAAGGGCCACAGAGCTTCTCTGCTGGCGAGGTGCACCC 1677
QY 529 SerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAlaArgAla 546
Db 1678 AGCACAGAGGCATCAGCCAGACAGGCGCCCGCCAGGAGCCCAACCACTCCCACTGCC 1731
```

RESULT 19

```
US-10-115-831-34
; Sequence 34, Application US/10115831
; Publication No. US20030219743A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219743A1el Nucleic Acids and
; FILE OF INVENTION: Polypeptides
; FILE REFERENCE: 792CIP2ADIV
; CURRENT APPLICATION NUMBER: US/10/115,831
; CURRENT FILING DATE: 2002-04-02
; PRIOR APPLICATION NUMBER: 09/667,298
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 34
; LENGTH: 2356
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (85)..(1554)
US-10-115-831-34
```

Alignment Scores:

Pred. No.: 3.3e-92 Length: 2356

```
Score: 1291.50 Matches: 285
Percent Similarity: 67.62% Conservative: 70
Best Local Similarity: 54.29% Mismatches: 124
Query Match: 46.37% Indels: 46
DB: 16 Gaps: 9
```

US-10-023-523-44 (1-546) x US-10-115-831-34 (1-2356)

```
QY 29 GlyAlaGlnGluArgProSerGlnAlaLaProAlaValGluAlaGluGlyProGlySer 48
Db 10 GCGCGCCGAGAG-----GCGACTGAGCGCGGAGCGGCGGCGGACGG 48
QY 49 SerGlnAlaProArgLys-----ProGluGlyAlaGlnAlaArgThrAlaGlnSerGly 66
Db 49 CGACGAGCGCCCGCGAGAGTTGAAATGGCACAATGGAAGAGAGCTGGAATTTGTGGG 108
QY 67 AlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyr 86
Db 109 CTAGGGGTGAAAGCAGATATGTTGTAACTCTCAATCAATGATATCTTCTCAACATCAA 168
QY 87 CysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGlu 106
Db 169 -----GGCTCAAAATTTGGTGGCACAAGTAAACAGCATTCATTGGAAGAGGATGAAGGC 222
QY 107 ProGluAspAlaGluLysSerArgThrTyrValAlaArgAsnGlyGluProGluProThr 126
Db 223 AGTGACTTTATAACACAGACAGAGGAATTTGGTGAC-----CCAGCA 264
QY 127 ProValValAsnGlyGluLysGluProSerLysGlyAspProAsnThrGluGluLeuArg 146
Db 265 TACTCAGCCAGAAATCAAGAGAGGAATCCCTGGGGGA-----GAAGCTCGA 312
QY 147 GlnSerAspGluValGlyAspArgAspHisArgArgProGlnGluLysLysLysAlaLys 166
Db 313 ACAGATCCCTCTGATGGTGCAGAGATTCAGAGTCCACAGACAGAAAGAAACT--- 369
QY 167 GlyLeuGlyLysGluLeuThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGlu 186
Db 370 ---TTAGGAAAGAAAGTTTATTACTGATGCAAGCCCTAAACACCCCTTTCAACCCACAG 426
QY 187 GlyLysLeuAlaAlaLeuCysLysLysTyrAlaGluLeuLeuGluGluHisArgAsnSer 206
Db 427 GAGAGCTGGCAGCTCTCTGTAAGAAATATGCTGATCTCTGGAGGAGAGCAGAGGTGTT 486
QY 207 GlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGlnLysAspHis 226
Db 487 CAGAAGCAAATGAAGATCTCTCAGAGAAGCAAGCAAGCCAGATTGTGAAGAGAAAGTTTCA 546
QY 227 LeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArg 246
Db 547 TTGCAGAGTGAACATAGCAGGCTATCTTGGCAAGAGCAAGCTAGAAATCTCTTTGCAGA 606
QY 247 GluLeuGlnArgHisAsnArgSerLeuLysGluGlyValGlnArgAlaArgGluGlu 266
Db 607 GAATTCAGCGTCACAATAGACGTTAAGAGGAGAAATATGCAAGCAGGACAGAGAGAA 666
QY 267 GluGlnLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeu 286
Db 667 GAAGAAGCAGCTAAAGAGCAACTGCACTTCCAGATTACCTTTAAATGAAATCAAGGCC 726
QY 287 GlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGluLeuAla 306
Db 727 CAGCTGGAGCAGCATGACATCCACAGCCAACTCCAGAGGAAACATTTAGCTGGGG 786
QY 307 GluArgLeuLysLysLeuLeuGluGlnTyrGluLeuArgGluGluHisIleAspLysVal 326
Db 787 GAGAAGCTAAAGAGCTCATCGAAGCATGACGCTGAGGAGAGAGCAGCATTTGATAGGTG 846
QY 327 PheLysHisLysAspLeuGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGlu 346
Db 847 TTCAACATAGGAACCTGCAACAGCAGCTCGTGATGCCAAACTGCAGCAAAACGACACAA 906
QY 347 MetLeuLysGluAlaGluGluArgHisGlnArgGluLysAspPheLeuLeuLysGluAla 366
```

Db 907 CTGATAAAGAGCTGATGAAAAACATCAGAGAGAGAGAGTATTATTAAAGAGCG 966
 QY 367 valGluSerGlnArgMetCysGluLeuMetLysGlnGlnThrHisLeuLysGlnGln 386
 Db 967 ACAGAAATCAGGACCAAAATACAAATGAAATGAAACAGCAAGTACAACTAAACACGAG 1026
 QY 387 LeuAlaLeuThrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGlu 406
 Db 1027 CTTTCTTTATATGATAAGTTTGAAGATTCAGACTACCAATGCGCAAAAGCAATGAA 1086
 QY 407 valPheThrThrPheLysGlnGlnMetGluLysMetThrLysLysLysLysLysLeuGlu 426
 Db 1087 CTGTTTACACCTTCAGACAGGAAATGGAAGATGACAAAGAAATTTAAAAACCTGGAA 1146
 QY 427 LysGluThrThrMetTyArgSerArgThrGluSerSerAsnLysAlaLeuLeuMet 446
 Db 1147 AAAGAAACAAATAATTGGCGTACCAATGGAAGAAACATAATAAGCACCTTCGCAATG 1206
 QY 447 AlaGluGluLysThrValArgAspLysGluLeuGluGluLysLysLysLysLysGlnArg 466
 Db 1207 GCTGAAGAGAAACAGTCCGTGATAAGAGTACAAAGGCCCTTCAAATATAAACTGGAACGG 1266
 QY 467 LeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgValGln 486
 Db 1267 TTAGAGAGCTGTGAGGCTCTTCAACACAGAAAGGATGAGCTCANTGAGAGGTGAA 1326
 QY 487 -----AspLeuSer 489
 Db 1327 GTCCTGAAAGACAGGTATCCATCAAGCGCCATCAAGCGCGCAACAGGGAATTAGCA 1386
 QY 490 AlaGlyGlyGlnGlySerLeuThr-----AspSerGlyProGluArgArgProGluGly 507
 Db 1387 ACACCTGTGATCCAGCCCTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1446
 QY 508 Pro-----GlyAlaGlnAlaProSer-SerProArgValThrGluAlaProCysTy 524
 Db 1447 AAAAGAGCCCTGGAGCGCACCTGGAGGCTGAGGCCCAAGATCAGAGAGCGCTGTGCA 1506
 QY 524 rProGlyAlaPro 528
 Db 1507 AGCCCGCTGCTCA 1519

RESULT 20

US-09-960-253-172
 ; Sequence 172, Application US/09960253
 ; Patent No. US20020123619A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Benson, Darin R.
 ; APPLICANT: Mohamath, Raedoh
 ; APPLICANT: Lodes, Michael J.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.556
 ; CURRENT APPLICATION NUMBER: US/09/960/253
 ; CURRENT FILING DATE: 2001-09-20
 ; NUMBER OF SEQ ID NOS: 187
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 172
 ; LENGTH: 2045
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-960-253-172

Alignment Scores:

Pred. No.:	7,58e-92	Length:	2045
Score:	1286.00	Matches:	284
Percent Similarity:	67.45%	Conservative:	71
Best Local Similarity:	53.99%	Mismatches:	132
Query Match:	46.18%	Indels:	39
DB:	9	Gaps:	8

US-10-023-523-44 (1-546) x US-09-960-253-172 (1-2045)

QY 28 GluGlyAlaGlnGluArgProSerGlnAlaAlaProAlaValGlnAlaGluGlyProGly 47
 Db 70 GAGGACAGCGGGGAAAGAGCG 129
 QY 48 SerSerGlnAlaProAlaGlyLys-----ProGluGlyAlaGlnAlaArgThrAlaGlnSer 65
 Db 130 CGGCGACG 189
 QY 66 GlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGluLeuGluAspLeuSerThr 85
 Db 190 GGGCTAGGGGTGAAGACAGATATGTTGTAACTCTCAATCAATCAATCAATCAATCTTCAACAT 249
 QY 86 TyrCysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAlaGlnGlyGluProAla 105
 Db 250 CAA-----GGCTCAAAATTTGGTGGCACAAGTAAACAAGCATTCTTGAAGAGGATGAA 303
 QY 106 GluProGluAspAlaGluLysSerArgThrTyValAlaArgAsnGlyGluProGluPro 125
 Db 304 GGCAGTCTGACTTTTAAACAGAGAACAGGAATTTGGTGAGC-----CCA 345
 QY 126 ThrProValValAsnGlyGluLysGluProSerLysGlyAspProAsnThrGluGluLeu 145
 Db 346 GCATACTGCACGCAAGAAATCAAGAGAGGAAATCCCTGGGGGA-----GAAGCT 393
 QY 146 ArgGlnSerAspGluValGlyAspArgAspHisArgArgProGlnGluLysLysLysAla 165
 Db 394 CGAACAGATCCCTGATGTTGGTTCAGCAAGATTCAGATGCAACAGGAAACAAAGAAAAA 453
 QY 166 LysGlyLeuGlyLysGluLeuThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrPro 185
 Db 454 -----TTAGGAAAGAGTATTATTACTGATCAAGCCCTAAACACCTTTCAACCCCA 507
 QY 186 GluGluLysLeuAlaLeuLysLysLysTyAlaGluLeuLeuGluHisArgAsn 205
 Db 508 GAGGAGAGCTGGGAGCTCTCTGTAAAGAAATATGCTGATCTCTGGAGGAGAGCAGAGT 567
 QY 206 SerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAsp 225
 Db 568 GTTCAGAGCAATGAAGATCTTCAGAGCAAGCAAGCCAGATTGTGAAGAGAAAGT 627
 QY 226 HisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuLeuSerLysCys 245
 Db 628 CACTGCGAGGTGAACATAGCAAGCTTCTTGGCAGAGCAAGCAAGCTGAAATCTCTTTCG 687
 QY 246 ArgGluLeuGlnArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGlu 265
 Db 688 AGAGAACTTCAGCGCTCAACAATAAGACGTTAAAGGAGGAAATATGCGAGCAGCAGAGAG 747
 QY 266 GluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspLeuGln 285
 Db 748 GAAGAGAACGACGTAAGAGCAACTGCACATTTCAGATTACCTTAGATGAATTCAA 807
 QY 286 LeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGluLeu 305
 Db 808 GCCCAGCTGGAGCAGCATGATCCACACCGCAAACTCCGACAGGAAACAACTTGAGCTG 867
 QY 306 AlaGluArgLeuLysLysLeuLeuGluGlnTyGluLeuArgGluHisLysLys 325
 Db 868 GGGGAGAGGCTTAAAGAGCTTCAGCAAGCTAGCGCACTGAGGAGAGGACATTTATTAAG 927
 QY 326 ValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGln 345
 Db 928 GTGTTCAACGTAAGAACTGCAACAGCAGCTCGTGTGATGCCAACTGCGAGCAACAGCA 987
 QY 346 GluMetLeuLysGluAlaGluArgHisGlnArgGluLysAspPheLeuLysGlu 365
 Db 988 CAACTGATAAAGAGCTGATGAAACATCAGAGAGAGAGAGTATTATTATAAGAA 1047
 QY 366 AlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlnThrHisLeuLysGln 385
 Db 1048 GCGACAGATCGAGGACAAATACGAAACAAATGAAACAGCAGGAAAGTACAACTAAACAG 1107

Qy 386 GlnLeuAlaLeuThrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSer 405
Db 1108 CAGCTTCTCTTATATGGATAAGTTTGAGAGATTCAGACTACCATGCCAAAAGCAAT 1167
Qy 406 GluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLeu 425
Db 1168 GAACCTGTTTCAACCTTCAGACAGGAATGGAAGATGACAAAGAAAATTAAAAAAGCTG 1227
Qy 426 GluLysGluThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeuGlu 445
Db 1228 GAAAAAGAAACATAATTTGGGTACCAATATGGGAAACACATAATTAAGCACTTCTGCAA 1287
Qy 446 MetAlaGluLysThrValArgAspLysGluLeuGluGluLysLysLysLysLysLys 465
Db 1288 ATGGCTGAAGAGAAAACAGTCCGTGATAAAGAGTACAAAGCCCTTCAAAATAAAAGCTGAA 1347
Qy 466 ArgLeuLysLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgVal 485
Db 1348 CGGTAGAGAGCTGTGACGGCTCTTCAACAGAAAGGAATGAGCTCAATGAGAGGTG 1407
Qy 486 Gln-----AspLeu 488
Db 1408 GAAGTCCTGAAGAGCAGGTATCCATCAAGCGGCCCATCAAGCGGCAACAGGGATTTA 1467
Qy 489 SerAlaGlyGlyGlnGlySerLeuThr-----AspSerGlyProGluArgArgProGlu 506
Db 1468 GCAACACCTGTGATGACAGCCCTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1527
Qy 507 GlyPro-----GlyAlaGlnAlaProSer-SerProArgValThrGluAlaProCys 523
Db 1528 TCGAAAGAGCCCTGGAGCGGCACCTGGAGCTGAGCCCAAGAGTACAGAGAGCGCTGTG 1587
Qy 523 sTyrProGlyAlaPro 528
Db 1588 CAAAGGCCCGCTGCA 1603

RESULT 21

US-10-108-260A-1586
; Sequence 1586, Application US/10108260A
; Publication No. US2004005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US2004005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT FILING DATE: 2002-03-27
; CURRENT APPLICATION NUMBER: US/10108,260A
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1586
; LENGTH: 2194
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-1586

Alignment Scores:
Pred. No.: 9,89e-62 Length: 2194
Score: 903.50 Matches: 257
Percent Similarity: 47.28% Conservative: 38
Best Local Similarity: 41.19% Mismatches: 112
Query Match: 32.44% Indels: 217
DB: 16 Gaps: 18

US-10-023-523-44 (1-546) x US-10-108-260A-1586 (1-2194)

Qy 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerSerPro 20
Db 66 ATGAAGAACCAAGACAAAAGACGGGGCTGCCAAACATCCAAATCCAAAAGCAGCCCA 125
Qy 21 GlyGlnProGluAlaGlyProGluGluGluAlaGlnGluArgProSerGlnAlaAlaProAla 40
Db 126 GGACACCGGAAGCAGGACCCGAGGGAGCCCGAGGAGCCGAGGAGCCGAGGAGCCGAGGAG 185
Qy 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60

Db 186 GTAGAAGCAGGAGGTCCCGGCGAGCAGCCAGGCTCTCCGAAAGCGGAGGGGCTCAAGCC 245
Qy 61 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
Db 246 AGAAGCGGTCACTGCGGGCCCTTCGTGATGTCCTGAGGAGCTGAGCCGCAACTGAA 305
Qy 81 AspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyProGlyGluAspGlyAla 100
Db 306 GACATACTGAGCACATCTGTGTGACAAATAACAGGGGGGCCCCGGGAGGATGGGCA 365
Qy 101 GlnGlyGluProAlaGluProGluAspAla-----Glu-LysSerArgTh 115
Db 366 CAGGCTGAGCGGCTGAAACCCGAGATCTGCCCTTCTAGGTTCATPAGCCCAAGGTGA 425
Qy 115 rTyrValAlaArg-----AsnGlyGluProGluProThrProValValAsn----- 130
Db 426 TCGTGTGTGAGGAGCTAGAGGGCGCTCCCTGAGCCACCCCTTTCCTTGGCATACTTCA 485
Qy 131 -----GlyGluLys-----GluProSerLysGlyAspProAsnThrGln 143
Db 486 TCCTCTGGGAAACAAAGCTGCTTGTGTTTGGGGAGTTGGTTTGGTTTCTTATCCCTC 545
Qy 143 uGluIleArgGlnSerAspGluValGlyAspArgAspHisArgArgProGlnGlu----- 161
Db 546 AGCGCTGAGCATAGAGGCTTCTGGCCACTACAGTGAGACACAGCACTTCAGAAATCTG 605
Qy 162 -----LysLysLysAlaLysGlyLeuGlyLysGlyLysGlyLysGlyLysLeuLeu 175
Db 606 AATACCCCGTCTTCTCTCCCGCCGCAAGCAAAA-----AAGGACTTAGTACTACC 656
Qy 175 uMetGln-----ThrLeuAsnThrLeuSerThrProGln 186
Db 657 TGTGAGAGAGAGGTGCGAGGACTACAGCGCCCTGCTGCTTTCATTTACAGCCCTCCCA 716
Qy 186 uGluLysLeuAlaAlaLeuCysLysLysTyrAlaGluLeuLeuGluGluHis----- 203
Db 717 GACAGACAGGACCCCTCATCATCAACCAACTGAGCTTACCTGTAGGACCTTCCCTT 776
Qy 204 -----ArgAsn----- 205
Db 777 CCCCATCCAAAATAATGGAGTTATTTCCCTTATTTTCAGCAAGTCCAGTTGATTTTACCT 836
Qy 206 -----SerGlnLysGlnMetLysLe 212
Db 837 TTGAAGTAGCAGCTGAGTCTTACCTTCTCTCCATCCCTTCTCTCACCTGACACAGG 896
Qy 212 uLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHis 232
Db 897 TCTGAGCGCTCCTCTAGT-----AGGAGGACAGCCATCTCTTGGGATGCACAT 947
Qy 232 rLysAlaValLeuAlaArgSerLysLeuLeuSerLeuCysArg----- 246
Db 948 GTCTAGTCTTTCGCTTA-----GATATGGCAAGTCTTTGCCAACTGAGTAGGCTGTAT 1001
Qy 247 -----Gln 247
Db 1002 GTTCTTAGAGGATGTTTTCGCCCATTTCCCATTTACAGAGAATTCAGGACACAGA 1061
Qy 247 uLeuGlnArgHisAsnArgSerLeuLysGluGluGlyValGln----- 261
Db 1062 AGTGAGGCTTCCAGCCCATAGGTGATCAATCTGGGGTTCAGAGATTTGAGTGTGTTAT 1121
Qy 262 -----ArgAlaArgGln 265
Db 1122 TCGTTGCCTTCTTGGGAGCAGATTCCATCCATAAACCATGCTTACCAAGTCTGACTC 1181
Qy 265 uGluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGln 285
Db 1182 ACTGGAGAGAAACACGCTGAGGTTGGA-----AAGCTGACCTTCCAGAGACTTGG 1232
Qy 285 nLeuGlnMetGluGlnHis-----AsnGluArgAsnSerLysLeuArg 299

Db 1233 GGCCCATGTTGTGTGTACACATGGAGTCCATCATATCAGATTGAGATGGGGGCTGGG 1292
QY 299 gGlnGluAsnMetGluLeuAla-----GluArgLeuLysLysLe 312
Db 1293 CAAGTGGCCCTGTGTGTGTGGGCTACCTGTGAGAAAGGAGCGCTTGACAAAGCCG 1352
QY 312 uileGluGlnTyrGluLeuArgGluGluHisileAspLysValPheLysHisLysAspLe 332
Db 1353 ACTGCTCCACCATCTTTGTTG---CAGCATATCGACAAAGTCTTCAACACAAAGGACCT 1409
QY 332 uGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGln 352
Db 1410 ACAACAGCAGCTGTGTGTGATGCAAGCTCCAGCAGCCCGAGAGATGTAAGAGGCGAGA 1469
QY 352 uGluArgHisGlnArgGluLysAspPhe----- 361
Db 1470 AGACGGGACACAGCGGAGAGAGATTGTGTAGGCTCAGGCCCCAGGGTTGGGGTGGGG 1529
QY 361 ----- 361
Db 1530 TGGGAGGAGACAGGCTGGGCTCTGGCTCAGCTCATAGCCGGTTATATGGGAGAGTCTG 1589
QY 361 ----- 361
Db 1590 GCCAGACGAGCAGATCTCCTGAGTACAGTCTGAGAGCAGGAGCCCTCAGTGGGTCT 1649
QY 361 ----- 361
Db 1650 GGTGCTTGTGCTAAACCAACATAGCCCTCGGGGCTTCTGACAGGATCTGGGGTTC 1709
QY 362 -----LeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetL 377
Db 1710 TGCTTGGAAATAGCTCTCGTAAAGAGGAGGAGTATGAGTCTGAGGATGTGAGCTGATGA 1769
QY 377 ysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluP 397
Db 1770 AGCAGCAGAGACCCACCTGAAGCAACAGCTTGCCTATATACAGAGAGATTGAGGAGT 1829
QY 397 heGlnAsnThrLeuSerLysSerSerGluValPheThrPheLysGlnGluMetGluL 417
Db 1830 TCCAGAACACACTTTCACAAAGCAGCAGGATTCACACATTCACAGCAGGAGTGA 1889
QY 417 ysMetThr 419
Db 1890 AGGTAAC 1897

RESULT 22

US-10-040-739-869

; Sequence 869, Application US/10040739

; Publication No. US20020173635A1

; GENERAL INFORMATION:

; APPLICANT: Jacobs, Kenneth

; McCoy, John

; Lavallie, Edward

; Racie, Lisa

; Merberg, David

; Treacy, Maurice

; Spaulding, Vikki

; TITLE OF INVENTION: SECRETED, EXPRESSED SEQUENCE TAGS

; NUMBER OF SEQUENCES: 1519

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genetics Institute, Inc.

; STREET: 87 Cambridgepark Drive

; CITY: Cambridge

; STATE: Massachusetts

; COUNTRY: U.S.A

; ZIP: 02140

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy Disk

; COMPUTER: IBM PC Compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/040,739
; FILING DATE: 07-Jan-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/036,520
; FILING DATE: 03-JUN-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 878-5851
; INFORMATION FOR SEQ ID NO: 869:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 529 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 869:
US-10-040-739-869

Alignment Scores:

Pred. No.: 1,05e-54 Length: 529
Score: 805.00 Matches: 154
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.35% Mismatches: 0
Query Match: 28.90% Indels: 0
DB: 14 Gaps: 0

US-10-023-523-44 (1-546) x US-10-040-739-869 (1-529)

QY 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerSerPro 20
Db 56 ATGAAGAACCAAGACAAAAGAACCGGCTGCCAAACAATCCAAATCCAAAAGAGAGCCCA 115
QY 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
Db 116 GACAAACCGAGCAGGAGCCCGAGGAGCCCGAGGAGCCCGAGGAGCCCGAGGAGCCCG 175
QY 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAla 60
Db 176 GTAGAAGCAGAGAGTCCCGGCGAGCAGCAGGCTCCTCCGAAAGCCGAGGAGGCTCAAGCC 235
QY 61 ArgThrAlaGlnSerGlyValAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 80
Db 236 AGAACGGCTCAGTCTGGGGCCCTTCGTATGTCTCTGAGAGAGCTGAGCGCCCACTGGAA 295
QY 81 AspileLeuSerThrTyrCysValAspAsnAsnGlnGlyProGlyGluAspGlyAla 100
Db 296 GACATACCTGAGCACATACCTGTGTGACAAATAACCAAGGGGGGGCCCGGCGAGGATGGGCA 355
QY 101 GlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsn 120
Db 356 CAGGGTGAGCGCGCTGAACCCGAGATGACAGAGAGTCCCGAGCTTCTGTCAGAGGAT 415
QY 121 GlyGluProGluProThrProValValAsnGlnGlyGluLysGluProSerLysGlyAspPro 140
Db 416 GGGGAGCCTGACCAACTCCAGTAGTCAATGGAGAGAGAAACCTCCAGGGGGATCCA 475
QY 141 AsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAsp 155
Db 476 AACACAGAGAGATCCCGCAGAGTACGAGGTGCGAGAGCCGAGAA 520

RESULT 23

US-09-976-740-51

; Sequence 51, Application US/09976740

; Publication No. US20020194633A1

; GENERAL INFORMATION:

; APPLICANT: Lees, Ann M.

; APPLICANT: Lees, Robert S.

; APPLICANT: Law, Simon W.

; APPLICANT: Arjona, Anibal A.

Db 2700 TAACTAATATGTTCCGGTGGTAGTGAGAAATAGTGTGTCAATTTTCAAGTAAACAGACT 2759
Qy 56 ----- 56
Db 2760 TAAAGAGTTAGGCAACGATTACTATAATTTCTTGATTAAAGATGTTTCGAATCTAAA 2819
Qy 56 ----- 56
Db 2820 TTCTGACAGGAAGTACTAGTTTCTGCTGAATGATACCTCATTTCTTGCTTCTCAGTTTCCATAAA 2879
Qy 56 ----- 56
Db 2880 AAAAAAGTTAGCAACATTTAACTCAACTGATGAGTTTGGCTGGCCCTGAAAAATCCC 2939
Qy 57 -----Gly 57
Db 2940 AACCAAGTGGTATAATCGTCTTCTTCTCACTTACCCCTCATCTCTCCCTGTGTAGGG 2999
Qy 58 AlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGlnGluLeuSerArg 77
Db 3000 CT-CAAGCCAGAACGGCTAGTCTGGGGCCCTTCGTGATGTCCTCAGGAGCTGAGCCCG 3058
Qy 78 GlnLeuGluAspIleLeuSerThrTyCysValAspAsnGlnGlyGlyProGlyGlu 97
Db 3059 CAACTGGAGACATACCTGAGCACATCTGTGTGGACAATAACACAGGGGGCCCCGGCGAG 3118
Qy 98 AspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGlnLysSerArgThrTyVal 117
Db 3119 GATGGGGCAGAGGTGAGCGGCTGAACCGAAGATGCAGAGAAGTCCCGGACCTATGTG 3178
Qy 118 AlaArgAsnGlyGluProGluProThrProValIaAsnGlyGluLysGluProSerLys 137
Db 3179 GCAGAGNATGGGAGGCTGNACCACTCCATGAGTCAATGGAGAGAAGNACCTCCAG 3238
Qy 138 GlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHisArg 157
Db 3239 GGGGATCCAAACACAGAGAGATCCGGCAGATGACGAGTCCGACAGCCGAGACCATCGA 3298
Qy 158 ArgProGlnLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 177
Db 3299 AGCCACAGAGAGAAGAAAGCAAGGGTTTGGGTGACGACAGAGCGGCTCTTTGTGAA 3358
Qy 178 ThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysTyAla 197
Db 3359 -----GCTGTGAGGAGAGGGAGTTTGGACTTCAGCTTCTCTCGGCC 3400
Qy 198 GluLeuLeu-----GluGluHisArgAsnSerGlnLysGlnMetLysLeuLeu 213
Db 3401 AGCTTTCTTCGACGATTCAAGGAACCGTACTTCTCAGAGCAGC---AAGTCACTC 3457
Qy 214 GlnLysLysGlnSerGln 219
Db 3458 TAGTCTTAATCAAGCCAG 3475
RESULT 24
US-10-671-242-51
; Sequence 51, Application US/10671242
; Publication No. US20040040049A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/671,242
; CURRENT FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849

;; PRIOR FILING DATE: 2000-03-02
;; PRIOR APPLICATION NUMBER: US 08/979,608
;; PRIOR FILING DATE: 1997-11-26
;; PRIOR APPLICATION NUMBER: US 60/031,930
;; PRIOR FILING DATE: 1996-11-27
;; PRIOR APPLICATION NUMBER: US 60/048,547
;; PRIOR FILING DATE: 1997-06-03
;; NUMBER OF SEQ ID NOS: 53
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 51
;; LENGTH: 22255
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-671-242-51

Alignment Scores:
Pred. No.: 9,68e-48 Length: 22255
Score: 740.50 Matches: 182
Percent Similarity: 40.56% Conservative: 7
Best Local Similarity: 39.06% Mismatches: 23
Query Match: 26.59% Indels: 255
DB: 13 Gaps: 4

US-10-023-523-44 (1-546) x US-10-671-242-51 (1-22255)

```
QY 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerSerPro 20
Db 2100 ATGAAGAACCAAGACAAAGACAGCGGGCTGCCAACAATCCAAATCCAAAAGCAGCCCA 2159
QY 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
Db 2160 GGACAAACCGAAGCAGACCCAGAGAGAGCCAGAGCGGCCAGCCAGCGGGCTCTGCA 2219
QY 41 ValGluAlaGluGlyProGlySerGlnAlaProArgLysProGlu 56
Db 2220 GTAGAAGCAGAAGGTCCCGGCAGCAGCCAGCTCTCCGAGAGCGGAGGTGTGCCAG 2279
QY 56 ----- 56
Db 2280 CTCTCGGTGCCAGCGGCGAGGAGAGCTGTGGGGTGGGCTCGCTTCTGGACTTAC 2339
QY 56 ----- 56
Db 2340 AGGCCGAGCCAGGTGTCCGGAGGAGGAGATGTAGATGAGAGACAGTGTGGGGC 2399
QY 56 ----- 56
Db 2400 CGCGTCCCGCTCTCGCGAGTTGGCGAGCTGCCCTCTAAGCACAGGAACAG 2459
QY 56 ----- 56
Db 2460 AGTTCTGGAGAGAGCTCCGAGGATTAAAGTCAGTGGCAGGCCAACAGGACCCAGT 2519
QY 56 ----- 56
Db 2520 CAGGAATCCAGGTCCCGTTAGAAACACCTCAGCCACGAGCTAACTGCCCTTCTGT 2579
QY 56 ----- 56
Db 2580 TTGAGGCATTTCAGATGATCTGAATGCCAAGAAATGGTTTGTGGGGGGAAGGAGA 2639
QY 56 ----- 56
Db 2640 TGGACTAGAAGTTGCTCCGTGCCATCCCTGTGCTGATGCTTTACATCTTTATGATC 2699
QY 56 ----- 56
Db 2700 TAAACAATATGTCGGGTGGTAGTGAAGATAGTTGTGTCATTTTACAAGTAACAGACT 2759
QY 56 ----- 56
Db 2760 TAAAGAAGTTAGGCAACGATTACTATAATTTCTTGATTTAAAGATGTTTCGAATCTAAA 2819
```

```
QY 56 ----- 56
Db 2820 TTCTCAGAGAACTAGATTGCTGATGATACTCCATTCTTCTCAGTTTCCATAAA 2879
QY 56 ----- 56
Db 2880 AAAAAAAGTTAGGCAACATTTAACTCAACTGATGAGTTGGCTGGCCCTGAAAAATCCC 2939
QY 57 ----- Gly 57
Db 2940 AACCAAGTGGTATAATCGTCTTCTTCTCACTCTACCCCTCATCTCTCTGCTGAGGG 2999
QY 58 AlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluLeuSerArg 77
Db 3000 CT-CAAGCCAGACCGCTCAGTCTGGGGCCCTTCGTGATGTCCTGAGGAGCTGAGCCG 3058
QY 78 GlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGlyGlu 97
Db 3059 CAACTGGGAAGACATCTGAGGACACATCTGTGTGACANTAAACAGGGGGCCCGCGGAG 3118
QY 98 AspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrVal 117
Db 3119 GATGGGCACAGAGGTGAGCGGCTGAACCCCAAGATGCAGAGAGTCCCGACCTATGTG 3178
QY 118 AlaArgAsnGlyGluProGluProThrProValValAsnGlyGluLysGluProSerLys 137
Db 3179 GCAAGGAATGGGAGCCTGAACCACTCCAGTCTCAATGGAGAGAGAACCTCCCAAG 3238
QY 138 GlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHisArg 157
Db 3239 GGGGATCCAAACACAGAGAGATCCGCAGAGTCCAGAGTCCGAGACCGAGACCATCGA 3298
QY 158 ArgProGlnGluLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGln 177
Db 3299 AGGCACACAGGAGAGAAAAAACCAAGGTTGGGTGAGCAGAGGGCGGCTCTTTGTGAA 3358
QY 178 ThrLeuAsnThrLeuSerThrProGluLysLeuAlaAlaLeuCysLysLysTyrAla 197
Db 3359 -----GCTGGTCCAGGAGGAGGTTGGACTTGGCTTCTCTGGGCC 3400
QY 198 GluLeuLeu-----GluGluHisArgAsnSerGlnLysGlnMetLysLeuLeu 213
Db 3401 AGTCTGTTCTGCCAGGATTCAGAGAAACGCTTCTCTCAGAGCAGC---AAGTCACTC 3457
QY 214 GlnLysLysGlnSerGln 219
Db 3458 TAGTCTAATCAAAGCCAG 3475
```

RESULT 25

US-10-023-523-51
; Sequence 51, Application US/10023529
; Publication No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03


```
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 22255
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-023-529-51

Alignment Scores:
Pred. No.: 9,68e-48 Length: 22255
Score: 740.50 Matches: 182
Percent Similarity: 40.56% Conservative: 7
Best Local Similarity: 39.06% Mismatches: 23
Query Match: 26.59% Indels: 255
DB: 14 Gaps: 4

US-10-023-523-44 (1-546) x US-10-023-529-51 (1-22255)
QY 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerPro 20
DB 2100 ATGAAGAACCAAGACAAAAGAAAGCGGGCTGCCAATCCAAATCCAAAGACGCCA 2159
QY 21 GlyLysProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaProAla 40
DB 2160 GGACAAACCGAAGCAGACCCGAGGAGCCAGAGCGGCCAGCCAGCGCGCTCTGCA 2219
QY 41 ValGluAlaGlnGlyProGlySerSerGlnAlaProArgLysProGlu 56
DB 2220 GTAGAAGCAGAAAGTCCCGCAGACAGCCAGGCTCTCTCGAAGCCGAGGCTGTGCCAG 2279
QY 56 56 2280 CTCTGCGTCCAGCGGCGAGGAGGAGCTGTGGGTCGGCTCGCTTCGACTTAC 2339
QY 56 56 2340 AGGCCAGGCGCAGGTTGTCGGGAGGAGGAGATGTAGATGAGAGACAGTGTCTGGGGC 2399
QY 56 56 2400 CGGGTCCCCCTCGCTGCGGAGTGTGGGAGTGCGGCCCTCTAAGCACAGGAACAG 2459
QY 56 56 2460 AGTTCTGGAGAGAAGTCCCGAGGATTAAGTCAGGTGGCGAGCAAAACGAGCACCCAGT 2519
QY 56 56 2520 CAGGAATCCAGTCCGTTAGAAACACCTCAGCCACCAGCAGCTAACTGCCCTTCTGT 2579
QY 56 56 2580 TTGAGGCATTCTAGATGATCTGAATGGCAAGAAATGGGTTTGTGGGGGGAAGAGA 2639
QY 56 56 2640 TGGACTAGAAGTTGCTCCGTCATCCCTGTGCTGATGCTTATACATCTTTATGATC 2699
QY 56 56 2700 TAACAAATATGTCGGGTGTGTAGTGAATAATGTTGTGTCATTTTACAGTAACAGACT 2759
QY 56 56 2760 TAAAGAGTTAGCAACGATTACTATAATTTCTGATTAAAGATGTTTTCGAATCTAA 2819
QY 56 56 2820 TTCTGACAGGAACTAGATTGTCGAATGATGATCTCAATCTTCTGCTTCTCAGTTTCCATAA 2879
QY 56 56 2880 AAAAAAGTTAGCAACATTTAACTCAACATGATGAGTTTGGCTGGGCTGAAATAATCCC 2939
```

```
QY 57 -----Gly 57
DB 2940 AACCAAGTGTATAATCGTCTTTCTTCTCACTCTACCCCTCATCTCTCTCTCTCTAGGG 2999
QY 58 AlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArg 77
DB 3000 CT-CAAGCCAGAACGGCTCAGTCTGGGGCCCTTGTGATGCTCTGAGAGCTGAGCCGC 3058
QY 78 GlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyProGlyGlu 97
DB 3059 CAACTGGGAAGACATACTCAGCACATCTGTGTGGACAATAACAGGGGGGCCCCGGCAG 3118
QY 98 AspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrVal 117
DB 3119 GATGGGGCACAGGGTGAGCCGGCTGAACCCGAAGATGAGAGAAGTCCCGGACCTATGTG 3178
QY 118 AlaArgAsnGlyGluProGluProThrProValValAsnGlyGluLysGluProSerLys 137
DB 3179 GCAAGGAATGGGAGCCTGAACCAACTCCAGTAGTCAATGGAGAGAAGAACCCCTCCAAG 3238
QY 138 GlyAspProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAspHisArg 157
DB 3239 GGGGATCCAAACACAGAAAGATCCGGCAGAGTCCGAGGTCGAGAGCCGAGACATCGA 3298
QY 158 ArgProGlnGluLysLysAlaLysGlyLeuGlyLysGluLysGluLysLeuMetGln 177
DB 3299 AGGCCACAGGAGAAAGAAAAGCCAGGGTTGGGTGAGCAGAGGGCGGCTTTGTGCA 3358
QY 178 ThrLeuAsnThrLeuSerThrProGluLysLeuAlaLeuCysLysLysTyrAla 197
DB 3359 -----GCTGGTGAGGAGAGGAGTTTGGACTTGGACTTGGACTTCTCTGGGCC 3400
QY 198 GluLeuLeu-----GluGluHisArgAsnSerGlnLysGlnMetLysLeuLeu 213
DB 3401 AGTCGTCTGCCAGGATTCAGAGAAACGGTACTTCTCAGAGCAGC---AAGTCATCT 3457
QY 214 GlnLysLysGlnSerGln 219
DB 3458 TAGTCTAATCAAGCCAG 3475

RESULT 26
US-10-023-523-51
; Sequence 51, Application US/10023523
; Publication No. US20020152485A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 22255
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-023-523-51
```

Alignment Scores:

Pred. No.: 9,68e-48 Length: 22255
 Score: 740.50 Matches: 182
 Percent Similarity: 40.56% Conservative: 7
 Best Local Similarity: 39.06% Mismatches: 23
 Query Match: 26.59% Indels: 255
 DB: 14 Gaps: 4

US-10-023-523-44 (1-546) x US-10-023-523-51 (1-22255)

```

Qy 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerPro 20
Db 2100 ATGAAGAACCAAGCAAAAGACGGGCTGCCAAACATCCATCCAAAGACGCCA 2159
Qy 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
Db 2160 GGACACCGGAAGCAGGACCGGAGGAGCCAGGAGCGGCCAGCGGGCTCTGCA 2219
Qy 41 ValGluAlaGluGlyProGlySerSerClnAlaProArgLysProGlu 56
Db 2220 GTAGAAGCAGAAGTCCCGCAGCAGCCAGCTCTCGAAGCGGAGGTGTGCCAG 2279
Qy 56 56
Db 2280 CTCTGCGTTGCCAGCGGCGGAGGAGCTGTGGGTGCGCTCGCTTCGGACTTAC 2339
Qy 56 56
Db 2340 AGGCCAGGCGCAGGTTGTCCGGAGGAGGAGATGTAGAATGAGAGGACAGTGTGGGGC 2399
Qy 56 56
Db 2400 CGCGTCCCGCTCGCTGTGGGAGTTGGGAGCTGCCCTCTAAGCACAGGAACAG 2459
Qy 56 56
Db 2460 AGTTCGAGAGAGCTCCGACGGGATTAAGTCAGGTGGCAGCAACAGGACCCAGT 2519
Qy 56 56
Db 2520 CAGAAATCCAGGTCCCGTTAGAAACACCTCAGCCACCAGCAGTAAGTCCCTTCTGT 2579
Qy 56 56
Db 2580 TTGAGGCATTTCTAGATGATCTGAATGCGAAGAAATGGTTTTGTGGGGGGAAGAGA 2639
Qy 56 56
Db 2640 TGGACTAGAAGTTGTCCTGTCATCCCTGTCGTGTCGTGATCATCTTTATGATC 2699
Qy 56 56
Db 2700 TAACAATAATGTCGGGTGAGTGTGAGAAATAGTTGTGTGATTTTACAAGTAACAGACT 2759
Qy 56 56
Db 2760 TAAAGAAGTTAGCAACGATTAATAATTTCTTGAATTTAAAGATGTTTCGAATCTAA 2819
Qy 56 56
Db 2820 TTCTGACAGGAAGTAGATTTGCTGATGATGATCTCCATCTTGTCTCAGTTCCATAA 2879
Qy 56 56
Db 2880 AAAAAAGTTAGCAACATTTAACTCAACTGATGAGTTTGGTGGGCTGAAAAATCCC 2939
Qy 57 57
Db 2940 AACCAAGTGGTAAATCGTCTTTCTTCTCACTTACCCCTCATCTCTCGCTGTAGGGG 2999
Qy 58 AlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArg 77

```

```

Db 3000 CT-CAAGCCAGAACGGCTCAGTCTGGGGCCCTCGTGATGTCTCTGAGGAGCTGAGCCGC 3058
Qy 78 GlnLeuGluAspIleLeuSerThrTyrcysValAspAsnAsnGlnGlyProGlyGlu 97
Db 3059 CAACCTGGAAGACATACTGAGCACATCTGTGTGACATAAACACAGGGGGGCCCCGGCAG 3118
Qy 98 AspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyVal 117
Db 3119 GATGGGCGACAGGGTGAGCCGCTGAACCCGAAGATGACAGAAAGTCCCGACCTATGTG 3178
Qy 118 AlaArgAsnGlyGluProGluProThrProValValAsnGlyGluLysGluProSerLys 137
Db 3179 GCAAGGATGGGAGCCTGAACCACTCCAGTAGTCAATGAGAGAAAGAACCCCTCCAAG 3238
Qy 138 GlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHisArg 157
Db 3239 GGGGATCCAAACACACAGAGAGATCCGACAGATCCGAGGTCCGAGACCGAGACCATCGA 3298
Qy 158 ArgProGlnGluLysLysLysAlaLysGlyLysGluLysGluIleThrLeuLeuMetGln 177
Db 3299 AGGCCACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3358
Qy 178 ThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLalaLeuCysLysLysTyAla 197
Db 3359 -----GCTCGTGAGGAGGAGGAGTTTGACTTGACGTTCTCTGGGCC 3400
Qy 198 GluLeuLeu-----GluGluHisArgAsnSerGlnLysGlnMetLysLeuLeu 213
Db 3401 AGTCTGTCTCCAGGATTCAGAGAAACCGTACTTCTCAGAGCAGC---AAGTCACTC 3457
Qy 214 GlnLysLysGlnSerGln 219
Db 3458 TAGTCTAATCAAGCCAG 3475

```

RESULT 27

```

US-10-616-187-51
; Sequence 51, Application US/10616187
; Publication No. US20040013668A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/616,187
; CURRENT FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 22255
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-616-187-51

```

Alignment Scores:
 Pred. No.: 9,68e-48 Length: 22255
 Score: 740.50 Matches: 182
 Percent Similarity: 40.56% Conservative: 7
 Best Local Similarity: 39.06% Mismatches: 23
 Query Match: 26.59% Indels: 255

DB: 16 Gaps: 4
US-10-023-523-44 (1-546) x US-10-616-187-51 (1-22255)
QY 1 MetLysAsnGlnAspLysLysAsnGlyAlaAlaLysGlnSerAsnProLysSerSerPro 20
Db 2100 ATGAAGAACCAAGACAAAAGAACCGGGCTGCCAAACATCCATCCAAAAGACGCCCA 2159
QY 21 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 40
Db 2160 GGACAAACCGGAAGCAGGACCCGAGGAGCCAGGAGCGGCCAGCGGCTCTCTGCA 2219
QY 41 ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 56
Db 2220 GTAGACGAAGGTCCTCGGACAGCCAGCGCTCTCTCGAAGCGGAGGGTGTGTGCCAG 2279
QY 56 56
Db 2280 CTCTGCGTTCACAGCGGAGGAGGAGCTGTGGGTGCGGCTCGCTTCGACTTAC 2339
QY 56 56
Db 2340 AGGCCGAGGCCAGGTTGTCGGGAGGAGGAGATGTAGATGACGAGCAGTGTGGGGC 2399
QY 56 56
Db 2400 CGCGTCCCCCTGCGCTCTGGCGAGTTGGCGGAGCTGCCCTCTTAAGCACAGGAACAG 2459
QY 56 56
Db 2460 AGTCTGGAGAGAAGTCCGACGGGATTAAAGTCAGGTGGCAGCAAAACGAGCACCCACT 2519
QY 56 56
Db 2520 CAGGAATCCAGTCCCGTTAGAAACACCTCAGCCACGACGCTAACTGCCCTTCCGT 2579
QY 56 56
Db 2580 TTGAGCAATTCTAGATGATCTGAATGCGAAGATGGTTGTGGGGGGAAGGAGA 2639
QY 56 56
Db 2640 TGGACTAGAAGTTGCTCCGTCCATCCCTGTGTGCTGATCTTATACATCTTTATGATC 2699
QY 56 56
Db 2700 TAACAAATATGTCGGGTGTAGTAGAATAAGTTGTGTCTTACAAAGTAAACAGACT 2759
QY 56 56
Db 2760 TAAAGAAATTAGCAACGATTACTATAATTTCTTGATTTAAAGATGTTTTCGAATCTAA 2819
QY 56 56
Db 2820 TTCTGACGAACTAGATTGCTGAATGATCTCCATTTCTGCTTCCAGTTTCCATAA 2879
QY 56 56
Db 2880 AAAAAAGTTAGCAACATTTAACTCAACTGATGATTTGGCTGGGCTGAAATATCCC 2939
QY 57 57
Db 2940 AACCAAGTGTATAATGCTTCTTCTACTACTACCCCTCAPCCCTCTGCTGTAGGGG 2999
QY 58 AlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArg 77
Db 3000 CT-CAAGCCAGAACGGCTCAGTCTGGGCGCTTCTGTGATCTCTGAGAGCTGAGCCG 3058
QY 78 GlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyProGlyGlu 97
Db 3059 CAACTGGAAGACATACTGAGCACATACTGTGTGGACAATAACACAGGGGGCCCCGGCAG 3118
QY 98 AspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrVal 117

Db 3119 GATGGGCACACAGGTGAGCCGCTGAACCCGAGATGAGAGAGTCCCGACCTATG 3178
QY 118 AlaArgAsnGlyGluProGluProThrProValValAsnGlyGluLysGluProSerLys 137
Db 3179 GCAAGGAATGGGAGCCTGAACCAACTCCAGTAGTCAATGGAGAGAACCCCTCCAG 3238
QY 138 GlyAspProAsnThrGluGluLeuLeuArgGlnSerAspGluValGlyAspArgAspHisArg 157
Db 3239 GGGGATCCAAACACAGAGAGATCCGGCAGAGTACGAGGTGGAGAGCCGACCATCGA 3298
QY 158 ArgProGlnGlyLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGln 177
Db 3299 AGCCACAGGACAGAAAAAGCCAAAGGTTTGGGTGAGCAGAGGGCGCTCTTTGTGA 3358
QY 178 ThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaAlaLeuCysLysLysTyrAla 197
Db 3359 -----GCTGGTTCAGGAGAGGGAGTTTGGACTTTCACGTTCTCTGGGCC 3400
QY 198 GluLeuLeu-----GluGluHisArgAsnSerGlnLysGlnMetLysLeuLeu 213
Db 3401 AGTCTGTTCTGCCAGGATTCAAAGGAAACGTTACTTCTCAGACGAGC---AAGTCACCTC 3457
QY 214 GlnLysLysGlnSerGln 219
Db 3458 TAGTCTAATCAAGCCAG 3475
RESULT 28
US-09-833-381-1333
; Sequence 1333 Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: NO. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1333
; LENGTH: 614
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(614)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1333
Alignment Scores:
Pred. No.: 1,23e-42 Length: 614
Score: 652.50 Matches: 144
Percent Similarity: 83.66% Conservative: 25
Best Local Similarity: 71.29% Mismatches: 25
Query Match: 23.43% Indels: 8
DB: 9 Gaps: 1
US-10-023-523-44 (1-546) x US-09-833-381-1333 (1-614)
QY 166 LysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrPro 185
Db 9 AAAACTTTAGAAAGAGAGTTTATTACTGATGCAAGCCCTAAACACCTTTCAACCCCA 68
QY 186 GluGluLysLeuAlaLeuCysLysLysTyrAlaGluLeuLeuGluHisArgAsn 205
Db 69 CACCACAANTGGCAGCTCTCTGTAAGAAATATGCTGATCTTCTGGAGGAGACGAGT 128
QY 206 SerGlnLysGlnMetLysLeuLeuGlnLysGlnSerGlnLeuValGlnGluLysAsp 225
Db 129 GTTCAGAAAGCAATAGATCTCTGCAAGAGCAAGCCAGATTTGTGAAGAGAAAGATT 188

QY 226 HisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeu-Cy 245
 DB 189 CACTTGCAGAGTGAACATAGCAAGCTATCTTGGCAAGCAAGCTAGAAATCTCTTTG 248
 QY 245 sArgGluLeuGlnArgHisAsn--ArgSerLeuLysGluGluGlyValGlnArgAlaArg 264
 DB 249 CAGAGNACTTCAGCGCTCACAATTAGACCGTTAAAGGAGGAAATATGCGAGCANGCACNA 308
 QY 265 GluGluGluGlu-LysArgLysGluValThrSerHisPheGlnValThrLeuAsnAsp11 284
 DB 309 GAGGAGAGAGACCGCAACCAAGCAAGCAATTCAGATTCACCTTAATGAAT 368
 QY 284 eGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetG1 304
 DB 369 TCAAGCCAGCTGGAGCAGCATGACATCCACAGCGCAACTCCGACAGGAAACATTGA 428
 QY 304 uLeuAlaGluArgLeuLysLysLeuIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 324
 DB 429 GCTGGGGAGAGAGCTAAAGAAAGCTCATCGAACAGTACGCACTGAGGAGAGACACATTGA 488
 QY 324 pLysValPheLysHisLysAspLeuGlnGlnGln-LeuValAspAla---LysLeuGln- 342
 DB 489 TAAAGTGTTCAAACATAGGACTGCGACACACCTTGATGCGCCCAAACTGCAAN 548
 QY 343 GlnAlaGlnGluMetLeu-LysGluAlaGluGluArgHisGlnArgGluLys 359
 DB 549 CAAACGACACACTGATAAAAGGAAGCTGNTGAAAAACATCANAGAGAGAGA 600

RESULT 29

US-10-425-114-32431
 ; Sequence 32431, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovacic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(5313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 32431
 ; LENGTH: 1955
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: UC-ZMFLB73326B04_FLI
 US-10-425-114-32431

Alignment Scores:
 Pred. No.: 2,89e-30 Length: 1955
 Score: 501.00 Matches: 147
 Percent Similarity: 46.30% Conservative: 91
 Best Local Similarity: 28.60% Mismatches: 154
 Query Match: 17.99% Indels: 122
 DB: 13 Gaps: 17

US-10-023-523-44 (1-546) x US-10-425-114-32431 (1-1955)

QY 23 ProGluAlaGly-----ProGluGly-----AlaGlnGluArgProSerGln 36
 DB 263 CCGAGGCTAATCTCTCCCGCGGCTTCGTCGCCGACAGCGAGCTCAGACACGGAC 322
 QY 37 AlaAlaPro-----AlaValGluAlaGluGlyProGly 47
 DB 323 GCGGCTCTCCCTCCTCCGCGCCCATAGTTGACGATGCAATCGACTCCGATAGTCCGAC 382
 QY 48 SerSerGlnAlaProArgLysProGluGlyAlaGlnAlaArgThrAlaGlnSerGlyAla 67

DB 393 GCCACC-----AACCCCGGTGGTGAGAAACCCCTAAGTGATCCCTCCCTCC 430
 QY 68 LeuArgAspValSerGluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyCys 87
 DB 431 CGCTCCACCCGCCGAAGATGCTCTCAGCCGCCGCTGCTGAAGCATTCGACACACTTGT 490
 QY 88 ValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluPro 107
 DB 491 CTGGAC-----CTTGTTCGGATCCG 511
 QY 108 GluAspAlaGluLysSerArgThrTyValAlaArgAsnGlyGluProGluProThrPro 127
 DB 512 GAGCGCGCT-----CTAGGGGAGCAGCGACCCCACT 544
 QY 128 ValValAsnGlyGluLysGluProSerLysGlyAspProAsnThrGluGluLeuArgGln 147
 DB 545 GCT-----GCAAGAGGTGAAGGATCCTTGAGGAAATCGTGA 583
 QY 148 SerAspGluValGlyAspArgHisArgArgProGlnGluLysLysLysLysGly 167
 DB 584 TCAGAGCAAGTGGGC-----GCTCCAAGTGATCAAAAGGTGCTCAAGGG 628
 QY 168 LeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGlu 187
 DB 629 AGCGGTGAG-----CAAAACCC 646
 QY 188 LysLeuAlaAlaLeuCysLysLysTyValAlaGluLeuLeuGluHisArgAsnSerGln 207
 DB 647 AAGTTCATC-----AAGCATAGCAAGCTTGAG 673
 QY 208 LysGlnMetLysLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeu 227
 DB 674 AAGCAACAAGATTATTT-----CAACTAGCCAGCATACACAAAGTG 718
 QY 228 ArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGlu 247
 DB 719 GTTCAGAAAGGATCAAGCCATTGCGATCAAAAGATAGACTAGAAATCTCTTTGATAGGAA 778
 QY 248 LeuGlnArgHisArgSerLeuLysGluGlyValGlnArgAlaArgGluGlu 267
 DB 779 TTTGAGGTCAAAACAAATGCTAAAGGAGAGTGCAGGAGGTATCGACAGGAGCAAG 838
 QY 268 GluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGln 287
 DB 839 AACATGCGTACGGAATTTGCTGAAAAATTTGATCATGCCATAAAGGCTGCTAGTGCCAAA 898
 QY 288 Met---GluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGluLeuAla 306
 DB 899 CTTGAGGAGCAGAGAGTTGAGCCATTTCTCAGCTA---GAAGAGAACAAATACGTTGAGA 955
 QY 307 GluArgLeuLysLysLeuIleGluGlnTyArgGluLeuArgGluGluHisIleAspLysVal 326
 DB 956 AGTAAACTGAAAGACCTTGTGTGATCAATATATACATTACTCAGCAGAAATATGCTCACC 1015
 QY 327 PheLysHisLysAspLeuGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGlu 346
 DB 1016 TTGAAGAGAAATGCTGGAACTTGAACTTGCTGATCTGAGACTTCACACATCAAGAG 1075
 QY 347 MetLeuLysGluAlaGluArgHisGlnArgLysAspPheLeuLeuLysGluAla 366
 DB 1076 -----AAGGCTGCTCAGGAACATACACAAATGACG----- 1105
 QY 367 ValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlnThrHisLeuLysGlnGln 386
 DB 1106 TTGATGCTGAACAGTTCTCAGCTTATG---ACTACTGAGAAAGAACTCGGTGCGAA 1162
 QY 387 LeuAlaLeuTyThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGlu 406
 DB 1163 CTAGCTTCTCAGCGGAAAGATTTCAGCACTTTCAGGATGCTTGTCAAAAAGCAATGAA 1222
 QY 407 ValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLeuGlu 426
 DB 1223 GTCTTTGAAACTTACAGCAGGAGATGGAAGATGATTTTCAGTGTATTAAGATCTTAA 1282

Sequence 64592, Application US/10424599
 Publication No. US20040031072A1

GENERAL INFORMATION:

APPLICANT: La Rosa Thomas J
 APPLICANT: Kovalic David K
 APPLICANT: Zhou Yihua
 APPLICANT: Cao Yongwei
 TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53223)B

CURRENT APPLICATION NUMBER: US/10/424,599

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ IDS: 285684

SEQ ID NO 64592

LENGTH: 1696

TYPE: DNA

ORGANISM: Glycine max

FEATURE: Length: 1696

OTHER INFORMATION: Clone ID: PAT_MRT3847_29338C.1

US-10-424-599-64592

Alignment Scores:
 Pred. No.: 9-5e-27 Length: 1696
 Score: 457.00 Matches: 143
 Percent Similarity: 43.97% Conservative: 94
 Best Local Similarity: 26.53% Mismatches: 145
 Query Match: 16.41% Indels: 158
 DB: 13 Gaps: 17

US-10-023-523-44 (1-546) x US-10-424-599-64592 (1-1696)

Qy 1 MetLysAsnGlnAspLysLys-----AsnGlyAlaAla 11
 Db 181 ATGCAGAAATCCAGATGCAATCAACTCCCGAAGTGAATTCGTGCGAGATGGTTGTT 240
 Qy 12 LysGlnSerAsnProLysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGln 31
 Db 241 GAGAGTACTACAGAGCCCTCTGCTCTCTGCTGCACTCCA-----TCTCTGAA-----CAA 288
 Qy 32 GluArgProSerGlnAlaAlaProAlaValGluAlaGlyProGly-----47
 Db 289 GAGAAACCCCTGAAT-----AGTTACAGAGATGATGCTCCCTGACCTTATTCAT 339
 Qy 48 -----Ser 48
 Db 340 TCTAATGACGATCAATGATTTGGGAGAAAAAGAGTTTCAAAACATCATGACTGCTCA 399
 Qy 49 SerGlnAlaProArgLysProGluGlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeu 68
 Db 400 GCGGAGATCCCTCTAGAAGCTCTATGACGACGACAGACAGAGGGGCACAG-----447
 Qy 69 ArgAspValSerGluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrcysVal 88
 Db 448 -----ATTACCTCTCTGAGCTTAATCTGTTCCCGAGTCTCTGCTGCTCCAGTCTGCT 501
 Qy 89 AspAsnAsnGlnGlyProGlyGlu-----AspGlyAlaGlnGlyGluPro 104
 Db 502 GAGGTAATGAGCAACACACAGGAGAAATGCCAAAGTTTCAGATGGATCAACTGTTGACCT 561
 Qy 105 Alaglu-----ProGluAspAlaGluLysSer 113
 Db 562 TTGGAGACAAAGACCTGCTGTTAAAGACACTTGTCCACGAGACTGTTGATCTTCA 621
 Qy 114 ArgThrTyValAlaAlaArgAsnGlyGluProGluProThrProValValAsnGlyGluLys 133
 Db 622 -----AAAAACAAAAAACCCTGAAACTAGT-----645
 Qy 134 GluProSerLysGlyAspProAsnThrGluGluLeuArgGlnSerAspGluValGlyAsp 153
 Db 645 -----645
 Qy 154 ArgAspHisArgProGlnGlnLysLysLysAlaLysGlyLeuGlyLysGluLeuThr 173

Db 646 -----GAAAGCGCAAAAGTGCAAG-----666
 Qy 174 LeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCys 193
 Db 667 -----CGTACTCTTAATCA-----GAGAAGGAGCTTTTAGAATTACA 705
 Qy 194 LysLysTyAlaGluLeuLeuGluHisArgAsnSerGlnLysGlnMetLysLeuLeu 213
 Db 706 CTGAAGTACCAACAGTGTG-----726
 Qy 214 GlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLys 233
 Db 727 -----GCAGAAAGAGATGCA 741
 Qy 234 AlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArg 253
 Db 742 GCTCTTGGCGTCCGAGATAAGCTTGAATCACTGTGCGAGAGTTACACGCCCAATAAA 801
 Qy 254 SerLeuLysGluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluVal 273
 Db 802 ATGCTGATGGAAGAAATCAAAACGTGTATCACTCAGCGGCAAACTTAAGCGTGGACTG 861
 Qy 274 ThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGlu 293
 Db 862 TCAGCCCAAGTTTCAAGATGCAATCAAGATGCTCAGCAATAGCTTGAGGCGGAAGGAT 921
 Qy 294 ArgAsnSerLysLeuArgGlnGlnAsnMetGluLeuAlaGluArgLeuLysLysLeu 313
 Db 922 GATGC-CTTTCTCAGCTAAAGGAGATGACATGTTAAGAAACACAGCTTAAGACAGCTAGT 980
 Qy 314 GluGlnTyArgLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGln 333
 Db 981 GAACAATATGAATCACTCTGAGCAGCAACATGCACAGAGTTGAAGCAAAATCATCTGGAA 1040
 Qy 334 GlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnMetLeuLysGluAlaGluGlu 353
 Db 1041 CTACAGCTTGTGATTTAAATAATTAAACACATCAGGAAAAATTG-----GTTCAAGAA 1094
 Qy 354 ArgHisGlnArgGluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCys 373
 Db 1095 CGGTACACAGATAAAA-----GTATATGCAAGCAAGATATCT 1130
 Qy 374 GluLeuMetLysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyThrGluLys 393
 Db 1131 CAGTTATTAGCAACTGAGAGAGT---TTACGTTGCACTGCAACTGATGGAGAAAA 1187
 Qy 394 PheGluGluPheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGln 413
 Db 1188 TTCCAACAATCCAGGAAGCATTTGTTAAGAGCAATGAGATTTTGAACATTCACAA 1247
 Qy 414 GluMetGluLysMetThrLysLysLysLeuGluLysGluThrThrMetTyArg 433
 Db 1248 GAGATTGAGAGATGGCAAAATCAATCAAGAACTCAAGAAGGAGAAATCAATTTCTGAG 1307
 Qy 434 SerArgTyrGluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArg 453
 Db 1308 AGTAATCTGAGAGCTCTGATCTTACACTCATAGATTAGTTGATGAGCGGCGCATG 1367
 Qy 454 AspLysGluLeuGluGlyLeuGlnValLysGlnArgLeuGluLysLeuCysArgAla 473
 Db 1368 AAGAAACAACTGGAGAAAACTAAATAATCAGAAAGAAAACTGGAAATCATTTATCGCGTCA 1427
 Qy 474 LeuGlnThrGluArg-----AsnAspLeuAsnLysArgValGln 486
 Db 1428 CTTACGCGCAAGAAAGCAAAATTTCTCAGAGACAGCAATTAACCTCGTTCAA 1484

RESULT 33

US-10-425-114-6003
 ; Sequence 6003, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua

Db	538	----					----	TTGATGCTGTAACAAGTTTCTTCAGTGTATG----	ACTACTGAGAAG	579
Qy	382	HisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeu	401	----	----	TTGAGTTCAGTATTCAGGACATTCAGGATGCGCTTG	639	
Db	580	AACTCGGTGCAACTAGCTTCTACGGGAAAGATTTCAGCACATTCAGGATGCGCTTG	639	----	----	TTGAGTTCAGTATTCAGGACATTCAGGATGCGCTTG	639	
Qy	402	SerLysSerSerGluValPheThrThrPhelysGlnGluMetGluLysMetThrLysLys	421	----	----	TTGAGTTCAGTATTCAGGACATTCAGGATGCGCTTG	639	
Db	640	TCAAAAAGCAATGAAGTCTTTTGAACCTTACAAGCAGGAGATGGAAGAAGATGATTCAGTG	699	----	----	TTGAGTTCAGTATTCAGGACATTCAGGATGCGCTTG	639	
Qy	422	IleLysLysLeuGluLysGluThrThrMetTyrArgSerArgTrpGluSerSerAsnLys	441	----	----	TTGAGTTCAGTATTCAGGACATTCAGGATGCGCTTG	639	
Db	700	ATAAAGATCTTAAGAAGAGAACGAATTTCTGAGGGAAAAATGTGAGAACTCAGATATT	759	----	----	TTGAGTTCAGTATTCAGGACATTCAGGATGCGCTTG	639	
Qy	442	AlaLeuLeuGluMetAlaGluLysThrValArgAspLysGluLeuGluGlyLeuGln	461	----	----	TTGAGTTCAGTATTCAGGACATTCAGGATGCGCTTG	639	
Db	760	GCTATTGTGAAGCTCATTTGAAGAGCGTGAGCTAACAAAGAAAGCAAAATAGAGAAATTCAAA	919	----	----	TTGAGTTCAGTATTCAGGACATTCAGGATGCGCTTG	639	
Qy	462	ValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeu	481	----	----	TTGAGTTCAGTATTCAGGACATTCAGGATGCGCTTG	639	
Db	820	ANTCAAGGGAGGAAGCTCGAATCCCTGTGTGCAACACTACAGGCAGAAAGGAAA-----	873	----	----	TTGAGTTCAGTATTCAGGACATTCAGGATGCGCTTG	639	
Qy	482	AsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSerGlyPro	501	----	----	TTGAGTTCAGTATTCAGGACATTCAGGATGCGCTTG	639	
Db	873	-----	873	----	----	TTGAGTTCAGTATTCAGGACATTCAGGATGCGCTTG	639	
Qy	502	GluArgArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGlu	520	----	----	TTGAGTTCAGTATTCAGGACATTCAGGATGCGCTTG	639	
Db	874	-----CAAGGCCCTCCGCCAGTATTCAGATGCCCTTCTAGCCAAGAA	918	----	----	TTGAGTTCAGTATTCAGGACATTCAGGATGCGCTTG	639	
RESULT 34													
US-09-764-869-467													
; Sequence 467, Application US/09764869													
; Patent No. US20020061521A1													
; GENERAL INFORMATION:													
; APPLICANT: Rosen et al.													
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies													
; FILE REFERENCE: PC007													
; CURRENT APPLICATION NUMBER: US/09/764,869													
; CURRENT FILING DATE: 2001-01-17													
; Prior application data removed - refer to PALM or file wrapper													
; NUMBER OF SEQ ID NOS: 2442													
; SOFTWARE: Patent In Ver. 2.0													
; SEQ ID NO 467													
; LENGTH: 405.													
; TYPE: DNA													
; ORGANISM: Homo sapiens													
; FEATURE:													
; NAME/KEY: SITE													
; LOCATION: (377)													
; OTHER INFORMATION: n equals a,t,g, or c													
; NAME/KEY: SITE													
; LOCATION: (398)													
; OTHER INFORMATION: n equals a,t,g, or c													
US-09-764-869-467													
Alignment Scores:													
Pred. No.: 3,24e-19 Length: 405													
Score: 352.00 Matches: 83													
Percent Similarity: 80.00% Conservative: 5													
Best Local Similarity: 75.45% Mismatches: 16													
Query Match: 12.64% Indels: 6													
DB: 9 Gaps: 2													
US-10-023-523-44 (1-546) x US-09-764-869-467 (1-405)													
Qy	73	GluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnGln	92	----	----	TTGAGTTCAGTATTCAGGACATTCAGGATGCGCTTG	639	
Db	75	AAAGAACGGGCTGCCAACTGGAAGACATACCTGAGCACATCTGTGTGACACATTAACAG	134	----	----	TTGAGTTCAGTATTCAGGACATTCAGGATGCGCTTG	639	
Qy	93	GlyGlyProGlyGluAspGlyValGlnGlyCyluProAlaGluProGluAspAlaGluLys	112	----	----	TTGAGTTCAGTATTCAGGACATTCAGGATGCGCTTG	639	

Db 135 GGGGGCCCCGGGAGGATGGGGCCACAGGGTGAAGCGGCTGAACCCGGAAGATGCAGAGAAG 194
Qy 113 SerArgThrTyrValAlaArgAsnGlyGluPro-GluProThrProValValAsn---G1 131
Db 195 TCCCGGACCTATGTGGCAGAGATGGGAGGCTTGACCAACTTCCAKTAGTTCAATTGG 254
Qy 131 yGluGluSerProSerLysGlyAsp-ProAsnThrGluGluLeu-ArgGlnSer-AspG1 150
Db 255 AGAGAAGGAACCTCCAAAGGGGGATTCCAAACACAGAGAAGATTCCCGCAGAGTTGACGA 314
Qy 150 uValGlyAspArgAspHis---ArgArgProGlnGluLysLysAlaLysGlyLeuG1 169
Db 315 AGTTCCGAGCAGGAGGACCATTCGAAGGCCACAGAGAGAAAGCCAAAGGTTTITGGG 374
Qy 169 yLysGluIleThrLeuLeuMet 176
Db 375 GANGAGATCACGTTTGCTGATT 396
RESULT 35
US-10-091-504-467
; Sequence 467, Application US/10091504
; Publication No. US2003005908A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C1
; CURRENT APPLICATION NUMBER: US/10/091,504
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2442
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 467
; LENGTH: 405
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (377)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (398)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-091-504-467
Alignment Scores:
Pred. No.: 3,248-19 Length: 405
Score: 352.00 Matches: 83
Percent Similarity: 80.00% Conservative: 5
Best Local Similarity: 75.45% Mismatches: 16
Query Match: 12.64% Indels: 6
DB: 15 Gaps: 2
US-10-023-523-44 (1-546) x US-10-091-504-467 (1-405)
Qy 73 GluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnGln 92
Db 75 AAAGAACGGGGCTGCCAACTGGAAGACATCTAGACACATCTGTGTGCAATAAACCAG 134
Qy 93 GlyGlyProGlyGluAspGlyValAlaGlnGlyGluProAlaGluProGluAspAlaGluLys 112
Db 135 GGGGGCCCCGGGAGGATGGGGCCACAGGGTGAAGCGGCTGAACCCGGAAGATGCAGAGAAG 194
Qy 113 SerArgThrTyrValAlaArgAsnGlyGluPro-GluProThrProValValAsn---G1 131
Db 195 TCCCGGACCTATGTGGCAGAGATGGGAGGCTTGACCAACTTCCAKTAGTTCAATTGG 254
Qy 131 yGluGluSerProSerLysGlyAsp-ProAsnThrGluGluLeu-ArgGlnSer-AspG1 150
Db 255 AGAGAAGGAACCTCCAAAGGGGGATTCCAAACACAGAGAAGATTCCCGCAGAGTTGACGA 314
Qy 150 uValGlyAspArgAspHis---ArgArgProGlnGluLysLysAlaLysGlyLeuG1 169
Db 315 AGTTCCGAGCAGGAGGACCATTCGAAGGCCACAGAGAGAAAGCCAAAGGTTTITGGG 374

Qy 169 yLysGluIleThrLeuLeuMet 176
Db 375 GANGAGATCACGTTTGCTGATT 396
RESULT 36
US-10-227-577-467
; Sequence 467, Application US/10227577
; Publication No. US20040005575A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C2
; CURRENT APPLICATION NUMBER: US/10/227,577
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 10/091,504
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 09/764,869
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 467
; LENGTH: 405
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (377)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (398)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-227-577-467
Alignment Scores:
Pred. No.: 3,248-19 Length: 405
Score: 352.00 Matches: 83
Percent Similarity: 80.00% Conservative: 5
Best Local Similarity: 75.45% Mismatches: 16
Query Match: 12.64% Indels: 6
DB: 16 Gaps: 2
US-10-023-523-44 (1-546) x US-10-227-577-467 (1-405)
Qy 73 GluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnGln 92
Db 75 AAAGAACGGGGCTGCCAACTGGAAGACATCTAGACACATCTGTGTGCAATAAACCAG 134
Qy 93 GlyGlyProGlyGluAspGlyValAlaGlnGlyGluProAlaGluProGluAspAlaGluLys 112
Db 135 GGGGGCCCCGGGAGGATGGGGCCACAGGGTGAAGCGGCTGAACCCGGAAGATGCAGAGAAG 194
Qy 113 SerArgThrTyrValAlaArgAsnGlyGluPro-GluProThrProValValAsn---G1 131
Db 195 TCCCGGACCTATGTGGCAGAGATGGGAGGCTTGACCAACTTCCAKTAGTTCAATTGG 254

QY 131 yGlulysGluProSerLysGlyAsp-ProAenThrGluGluLeu-ArgGlnSer-AspG1 150
 Db 255 AGAGAGAGAACCTCCAGAGGGGATTCCAAACACAGAGAGATTCCGCGAGAGTTGACGA 314
 QY 150 uValGlyAspArgAspHis---ArgArgProGlnGluLysLysLysAlaLysGlyLeuG1 169
 Db 315 AGTTCGGAGCAAGCAACCATCGAAGGCCACAGGAGAGAAAGCCACAGGTTTGGG 374
 QY 169 yLysGluLeuThrLeuLeuMet 176
 Db 375 GAGGAGATCACGTTTGCTGATT 396

RESULT 37

US-10-264-049-1238
 ; Sequence 1238, Application US/10264049
 ; Publication No. US20040005579A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Birse et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: P133P1
 ; CURRENT APPLICATION NUMBER: US/10/264,049
 ; PRIOR FILING DATE: 2002-10-04
 ; PRIOR APPLICATION NUMBER: PCT/US01/18569
 ; PRIOR FILING DATE: 2001-06-07
 ; PRIOR APPLICATION NUMBER: US 60/209,467
 ; PRIOR FILING DATE: 2000-06-07
 ; NUMBER OF SEQ ID NOS: 4360
 ; SOFTWARE: PatentIn Ver. 3.1
 ; SEQ ID NO 1238
 ; LENGTH: 725
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (226)..(226)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (320)..(320)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (324)..(324)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (331)..(331)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (339)..(339)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (384)..(384)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (546)..(546)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (553)..(553)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (605)..(605)
 ; OTHER INFORMATION: n equals a,t,g, or c

; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (616)..(616)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (622)..(622)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (662)..(662)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (668)..(668)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (698)..(698)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (705)..(705)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (712)..(712)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; US-10-264-049-1238

Alignment Scores:

Pred. No.:	3.01e-16	Length:	725
Score:	318.00	Matches:	74
Percent Similarity:	77.05%	Conservative:	20
Best Local Similarity:	60.66%	Mismatches:	26
Query Match:	11.42%	Indels:	3
DB:	16	Gaps:	0

US-10-023-523-44 (1-546) X US-10-264-049-1238 (1-725)

QY	363	LeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHis	382
Db	334	ITAAANGAGGCGACAGAAATCGAGCACAATAACGACAAATGAACACACANGAAGTACAA	393
QY	383	LeuLysGlnGlnLeuAlaLeuTy-ThrGluLysPheGluGluPheGlnAsnThrLeuSer	402
Db	394	CTAAACACAGCAGCTTCCTTTATATGATAGTTTGAGATTTCCAGACTACCATGGCA	453
QY	403	LysSerSerGluValPheThrPhelLysGlnGluMetGluLysMetThrLysLysIle	422
Db	454	AAAAGCAATGAATGGTTTACAACTTCAGACAGGAAATGGAAGATGACAAAGAAATTT	513
QY	423	LysLysLeuGluLysGluThrThrMetTyArgSerArgTrpGluSerSerAsnLysAla	442
Db	514	AAAAACTGGAAAAGAAACCAATATTTGGCGNACCAANGGAAACCAT-AAATAAGCA	572
QY	443	LeuLeuGluMetAlaGluGluLysThrValArgAspLysGluLeuGluGluGlnVal	462
Db	573	CTTTGCAATGGCTCAAGAGGAAAAACAGTCCGNGATAAAGAGTNCAGGNCCTTCAATA	632
QY	463	LysIleGlnArg-LeuGluLysLeuCysArgAla-LeuGlnThrGluArgAsnAspLeuA	482
Db	633	AAACTGGAAACGGGTAGAGAACTTGGGCANGGCTTTTTCACACCGAAGGATGGGTCA	692
QY	482	sn 482	
Db	693	AT 694	

RESULT 38

US-09-876-143-1445
 ; Sequence 1445, Application US/09876143
 ; Publication No. US20040081958A1
 ; GENERAL INFORMATION:

; APPLICANT: Infogen Inc.
 ; APPLICANT: BILERTSEN, KENNETH J.
 ; APPLICANT: PFISTER-GENSKOW, MARTHA
 ; APPLICANT: CHILDS, LYNETTE
 ; APPLICANT: FORSYTHE, TODD
 ; APPLICANT: BISHOP, MICHAEL D.
 ; TITLE OF INVENTION: IDENTIFICATION AND USE OF MOLECULAR MARKERS INDICATING
 ; TITLE OF INVENTION: CELLULAR REPROGRAMMING
 ; FILE REFERENCE: 028040-0202
 ; CURRENT APPLICATION NUMBER: US/09/876,143
 ; CURRENT FILING DATE: 2001-06-06
 ; PRIOR APPLICATION NUMBER: 60/209,874
 ; PRIOR FILING DATE: 2000-06-07
 ; NUMBER OF SEQ ID NOS: 1744
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 1445
 ; LENGTH: 981
 ; TYPE: DNA
 ; ORGANISM: Bovine
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; LOCATION: (1)..(981)
 ; OTHER INFORMATION: n is a, c, g, or t
 US-09-876-143-1445

Alignment Scores:
 Pred. No.: 1,06e-15 Length: 981
 Score: 313.00 Matches: 74
 Percent Similarity: 62.88% Conservative: 9
 Best Local Similarity: 56.06% Mismatches: 46
 Query Match: 11.24% Indels: 4
 DB: 12 Gaps: 2
 US-10-023-523-44 (1-546) x US-09-876-143-1445 (1-981)
 QY 45 GlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAlaArgThrAlaGln 64
 DB 26 GTGGCGGGCGGCTCAGAGATAGTGGATCCCGCC---GGGTGCGAGGANTT-CGGCAGCAG 81
 QY 65 SerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGluAspIleLeuSer 84
 DB 82 GCTGAGCCCTCCGNGATGTTCTGAGGAGTTGANGCNCNCATGATGACATCTCTCANN 141
 QY 85 ThrTyrCysValAspAsnGlnGlyClyProGlyGluAspGlyAlaGlnGlyGluPro 104
 DB 142 ACATACCTCGTGAGCAACAGNANNNGGCGCCCGAGTGGATGTCGACNAGGNGAGCCT 201
 QY 105 AlaGluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsnGlyGluProGlu 124
 DB 202 GCTGAACCCGAGATGCANAGCNCNGGACCTATNCCTCANGANTGGNAGCCTGAG 261
 QY 125 Pro---ThrProValAsnGlyGlyGluProSerLysGlyAspProAsnThrGlu 143
 DB 262 CCAGATATCCAGTAGTCTNNNGAGAGAAANNCTCCAGGGGGGAGCCGGCCCGGTC 321
 QY 144 GluIleArg-GlnSerAspGluValGlyAspArgAspHisArgArgProGlnGluLysLy 163
 DB 322 TAGATCCTGTACACAGNATGAGTTGAGACGAGNACCACCCAAAGCCNCGGATANTAA 381
 QY 163 slyslalysGlyLeuGlyLysGluIleThrLeu 174
 DB 382 GAANNCTANGGTGGGAANGAAATPACCTTG 415

RESULT 39

US-10-425-114-18551
 ; Sequence 18551, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(5313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 18551
 ; LENGTH: 1035
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: LIB3088-030-A1_FLI
 US-10-425-114-18551

Alignment Scores:
 Pred. No.: 5.25e-15 Length: 1035
 Score: 304.50 Matches: 78
 Percent Similarity: 55.75% Conservative: 48
 Best Local Similarity: 34.51% Mismatches: 77
 Query Match: 10.93% Indels: 23
 DB: 13 Gaps: 6
 US-10-023-523-44 (1-546) x US-10-425-114-18551 (1-1035)
 QY 289 GluGlnHisAsnGluArgAsnSerLysLeuArgGlnGlnGlnGlnGlnGlnGlnGluArg 308
 DB 12 GAGCAGAGGGTTGAGTGCATTCCTCAGCTAGAGAGACATATG---TTGAGAGCTAAA 68
 QY 309 LeuLysLysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAspLysValPheLys 328
 DB 69 CTCAAAGACATTGCTGATCAGTATACATTACTCAGCAGAAATATGCTCACCATTGAAA 128
 QY 329 HisLysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnGlnGlnGlnGlnGln 348
 DB 129 GAGAAATGATGGAACTTGAGCTGCTGATCTGAGACTTCAACAAATCAAGAG--- 182
 QY 349 LysGluAlaGluGluArgHisGlnArgGluLysAspPheLeuLeuLysGluAlaValGlu 368
 DB 183 AAGGTGCTCAGCAACATACCCAAATGCAA-----TTGTAT 218
 QY 369 SerGlnArgMetCysGluLeuMetLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 388
 DB 219 GCCGAGCAAGTTTCTAGCTATG---ACTACTGAGNAGAACCTGGGTGCAACTAGCT 275
 QY 389 LeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluValPhe 408
 DB 276 TCGGATGGAGAAAGATTTCAGCACTTTTCAGGATGCCCTGTCAAAAAGCAATGAAGTCTT 335
 QY 409 ThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLysLysLysLysGlu 428
 DB 336 GAACCTTACAGCAGGAGATGGAAAATAATGATTTGGTGATAAAAATACTTAAGAGGAC 395
 QY 429 ThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeuGluMetAlaGlu 448
 DB 396 AATGAATTTCTGAAGGAAAAATGTGAGAACTCAGATATTGCTCTGTGAAGCTCATTTAA 455
 QY 449 GluLysThrValArgAspLysGluLeuGluGlyLeuGlnValLysIleGlnArgLeuGlu 468
 DB 456 GAGGTGAGCTTACGAGAGAGCAATAGAGAAATTTGAAAACCAAGAGGAGACACTTGGAG 515
 QY 469 LysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgValGlnAspLeu 488
 DB 516 TCCCTGTGTCGATCATTACAGGAGAAAGAAACAA----- 551
 QY 489 SerAlaGlyGlyGlnGlySerLeuThrAspSerGlyProGluArgArgProGluGlyPro 508
 DB 552 -----GGCGCTCCGCCAGTATTCAGACGCC---CCITCTAGCAAGAGACATGCCA 602
 QY 509 GlyAlaGlnAlaProSer 514
 DB 603 GCGACAAAGTCAAGAAATCT 620


```
Qy 305 -LeuAlaGluArgLeuLysLysLysLeuLeuLeu-----GlnTyrGluLeu----- 318
Db 4860 CCTGCAGCTCAGCAAGAGAGAGCTGGAAGGGGACCTGAAGACCTGAGGCTTCAGGC 4919
Qy 319 -----ArgGluGluHisIleAspLysValPheLys----- 328
Db 4920 CGACTCTGCCATCAAGGGGAGGAGAGCCATCAAGCAGCTACGCAACTGCAGGCTCA 4979
Qy 329 -HisLysAspLeuGlnGlnLeuValAspAlaLys----- 340
Db 4980 GATGAAGGACTTCAAGAGAGAGTGGAGATGCCGTCCTCCAGAGATGAGATCTTTGC 5039
Qy 341 -----LeuGlnGlnAlaGln 345
Db 5040 CACAGCCAAAGAGATGAGAGAGAGCAAGAGCTTGAAGACGACCTCATGACGCTACA 5099
Qy 345 nGluMetLeuLysGluAlaGluArgHisGlnArg-----GluLysAspPheLe 362
Db 5100 AGAGGACCTCCCGCGCTGAGAGGCTCGCAAAAGCGGACCTCGAAGAGGAGAACT 5159
Qy 362 uLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHi 382
Db 5160 GGCAGAGAGAGTGGCCAGTAGCTGTCCGGAAGGAACCCACTCCAGGACGAGAGCCCG 5219
Qy 382 sLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSe 402
Db 5220 CTGGAGCCCGGATCCGCTGAGGAGAGAGCTGGAGGAGGAGGAGGACGACATGGA 5279
Qy 402 rLysSerSerGluValPheThrThrPheLysGlnGlnMetGluLysMetThrLysLysI 422
Db 5280 GGCCATGAGGCAC-----CGGTCGCGCAAGAGCCACACAGCAGGC 5318
Qy 422 eLysLysLeuGluLysGluThrMetTyrArgSerArgTrpGluSerSerAsnLysAl 442
Db 5319 CGAGCAGTCAACACAGAGCTGGCCACAGAGCGGACGCGCCGCAAGAGATGAGATGC 5378
Qy 442 aLeuLeuGluMet---AlaGluGluLysThrValArgAspLys-----GluLeuGluG 459
Db 5379 CCGGACGAGCTCGAGCGGAGAGAGAGCTCGGAGCAAGCTCCAGGAGATGGAGGG 5438
Qy 459 Y-----LeuGlnValLysLleGlnArgLe 467
Db 5439 GGCCGCTCAAGTCAAGTTCAAGTCCACCATCCGCGCGCTGGAGGCCAAGATTGCACAGCT 5498
Qy 467 uGluLysLeu-----CysArgAlaLeuGlnThrGluArgAsnAspLeu----- 481
Db 5499 GGAGGAGCAGGTGCGAGGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5558
Qy 482 ---AsnLysArgValGlnAspLeuSerAlaGlyGlnGlnGlySerLeuThrAsp----- 498
Db 5559 GAAAGACAAGAGCTGAAGGAATCTT-GCTGCAGGTGGAGGAGGAGGAGGAGGAGGAG 5617
Qy 499 -----SerGlyProGluArgArgProGluGlyProGlyAlaGlnAlaProSe 514
Db 5618 AGCAGTACAAGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5677
Qy 514 rSerProArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAlaSerGl 534
Db 5678 TGGAGGAGGAGGAGGAGGAGTCCAGCGCATCAACCGCAACCGAGGAGAGTGCAGCGGG 5737
Qy 534 yGln-----ThrGlyPro-----GlnGln 540
Db 5738 AGCTGGATGAGGCCACGAGAGAGCAACGAGGAGGAGTGGCGCGGAGGTGAACGCTACA 5797
Qy 540 uProThrSerAlaArg 545
Db 5798 GCAAGCTCAGGCGAGG 5813
```

RESULT 42

US-10-171-311-161

; Sequence 161, Application US/10171311

; Publication No. US2003008720A1

GENERAL INFORMATION:

```
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoerish, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; PRIOR FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 161
; LENGTH: 6861
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-171-311-161
```

Alignment Scores:

```
Pred. No.: 1,95e-12 Length: 6861
Score: 284.00 Matches: 150
Percent Similarity: 37.69% Conservative: 101
Best Local Similarity: 22.52% Mismatches: 231
Query Match: 10,20% Indels: 185
DB: 15 Gaps: 25
```

US-10-023-523-44 (1-546) x US-10-171-311-161 (1-6861)

```
Qy 9 GlyAlaAlaLysGlnSerAsnProLysSerSerProGlyGlnProGluAlaGlyProGlu 28
Db 4051 GGAGCTGCTTCAAGAGAGAAACCGGAGAGAGCTCAACGCTCTACGAGCTCGCCAGCT 4110
Qy 29 GlyAlaGlnGluArgProSerGlnAlaAlaProAlaValGluAlaGluGlyProGlySer 48
Db 4111 GGA-----GGAGGA 4119
Qy 49 SerGlnAlaProArgLysProGluGly---AlaGlnAlaArgThrAlaGlnSerGlyAla 67
Db 4120 GCGGACAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4179
Qy 68 LeuArgAspValSerGluLeuLeuSerArgGlnLeuGluAspIleLeuSerThrTyCys 87
Db 4180 CCACATCTCCACTCTCAACATCCAGCTCTCCGACTCGAAGAGAGAGAGAGAGAGAGAG 4239
Qy 88 ValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluPro 107
Db 4240 CAGCAGCTGGAAGCTCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4299
Qy 108 -----GluAspAlaGluLys 112
Db 4300 CACCCAGCAGTACGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4359
Qy 113 SerArgThrTyrValAlaArgAsn-----GlyGluProGluProThrProVal 128
Db 4360 GCTTCAGCAGGAGCTGAGCAGCAGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 4419
Qy 129 ValAsnGlyGluLysGluPro-----SerLysGlyAspProAsnThr 142
Db 4420 CAACCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4479
Qy 143 GluGluIleArgGlnSerAspGluValGly-AspArgAspHisArgArgProGlnGlu 162
Db -----
```


QY 49 SerGlnAlaProArgLysProGluGly---AlaGlnAlaArgThrAlaGlnSerGlyAla 67
: : : : :
Db 4120 GCGGAACACCTGCAAGACCAAGCTGGACGAGATGGAGCCAGCAGAACTGGAGCG 4179
QY 68 LeuArgAspValSerGluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyCys 87
: : : : :
Db 4180 CCACATCTCCACTCTCAACATCCAGCTCCGATCGCAAGAGAAAGCTGGAGACTTTCG 4239
QY 88 ValAspAsnAsnGlnGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluPro 107
: : : : :
Db 4240 CAGCACCTGGAAGCTCTGGAAGAGGGAAGAGAGGTTCCAGAGGAGATCGAGAACT 4299
QY 108 -----GluAspAlaGluLys 112
Db 4300 CACCAGCAGTACGAGGAGAGAGGCGCGCTTATGATAAACTGGAAAGAGCAAGAAACAG 4359
QY 113 SerArgThrTyValAlaArgAsn-----GlyGluProGluProThrProVal 128
: : : : :
Db 4360 GCTTCAGCAGAGCTGGAGACCTGCTGTTGATTTGGCAACCAAGCGGCACTCTGTC 4419
QY 129 ValAsnGlyGluLysGluPro-----SerLysGlyAspProAsnThr 142
: : : : :
Db 4420 CAACCTGGAAGAGAGCAGAGGAAATTTGATCAGTTGTTAGCCGAGGAGAAACATCTC 4479
QY 143 GluGluIleArgGlnSerAspGluValGly-AspArgAspHisArgArgProGlnGluLys 162
: : : : :
Db 4480 TTCCAAATACGC-----GGATGAGAGGAGACAGACTGAGCGCAGAGCCAG 4524
QY 162 sLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThrLe 182
: : : : :
Db 4525 GGAGAG-----4531
QY 182 uSerThrProGluGluLysLeuAlaLeuCysLysLysTyAlaGluLeuLeuGluL 202
: : : : :
Db 4532 -----GAAACCAAGGCGCTGCTCCCTGGCTGGCGCTTGAAGCGCTTGAAGC 4581
QY 202 uHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValG 222
: : : : :
Db 4582 CAAGAGAACTCGAGCGGACCAACAAATGCTCAAGCCGAATGGAAGACCTGGTCAG 4641
QY 222 nGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuG 242
: : : : :
Db 4642 CTCCAAGATGACGTGGGCAAGACCTCATGAGCTGGAGAAGTCCAAGCGGCGCTTGA 4701
QY 242 uSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGluGlyValGlnAr 262
: : : : :
Db 4702 GACCCAGATGAGGAGATGAAGACGACGCTGGAAGAGCTGGAGACGAG-----CTGCA 4755
QY 262 gAlaArgGluGluGluLysArgLysGluValThr-----SerHisPh 277
: : : : :
Db 4756 AGCCAGGAGGACGCCCAACTGGGCTGAAGTCAACATGACGAGCGCTCAAGGGCCAGTT 4815
QY 277 eGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLy 297
: : : : :
Db 4816 CGAA-----AGGGATCTCCAAAGCCCGGAGCAGAGATGAGGAGAGAGAGGCA 4866
QY 297 sLeuArgGlnGlnAsnMetGlu-----304
: : : : :
Db 4867 ACTGCAGAGACAGCTTCACAGATGATGACAGCAACTGGAACGAGCGGAAGCAACGTGC 4926
QY 305 -LeuAlaGluArgLeuLysLysLeuIleGlu-----GlnTyGluLeu----- 318
: : : : :
Db 4927 CTTGGCAGCTGCAGCAAGAGAAAGCTGGAAGGGGACCTGAAGACCTGGAGCTTCAGGC 4986
QY 319 -----ArgGluGluHisLeuAspLysValPheLys-----328
: : : : :
Db 4987 CGACTCTGCCATCAAGGGGAGGAGAGAGCCATCAAGCAGCTACGCAACTTGCAGGCTCA 5046
QY 329 -HisLysAspLeuGlnGlnGlnLeuValAspAlaLys-----340
: : : : :
Db 5047 GATGAAGGACTTTCAAAGAGAGCTGGAAGATGCCCGTCCCTCCAGAGATGAGATCTTTC 5106
QY 341 -----LeuGlnGlnAlaG 345

Db 5107 CACAGCAAAAGAGAATGAGAAGAAACCAAGAGCTTGAAGCAGACCTCATGCAGCTACA 5166
QY 345 nGluMetLeuLysGluAlaGluArgHisGlnArg-----GluLysAspPheLe 362
: : : : :
Db 5167 AGAGGACCTCGCGCGCTGAGAGGGCTGCAAAACAAGCGGACCTCGAGAAGAGAACT 5226
QY 362 uLeuLysGluAlaValIleGluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHi 382
: : : : :
Db 5227 GGCAGAGGAGCTGGCCAGTAGCTGTGCGGAAGAACGCACTCCAGGACGAGAGCGCG 5286
QY 382 sLeuLysGlnGlnLeuAlaLeuTyThrGluLysPheGluGluPheGlnAsnThrLeuSe 402
: : : : :
Db 5287 CTTGGAGGCGCGGATCGCCAGCTGAGAGAGAGCTGGAGGAGGAGCGGCAACATGGA 5346
QY 402 rLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysI 422
: : : : :
Db 5347 GGCATGAGCGAC-----CGGTCGCAAAAGCCACACAGCAGCGC 5385
QY 422 eLysLysLeuGluLysGluThrThrMetTyArgSerArgTrpGluSerSerAsnLysAl 442
: : : : :
Db 5386 CGAGCAGCTCAGCAACGAGCTGGCCACAGAGCGCAGCAGCGGCCCAAGAATGAGAGTGC 5445
QY 442 aLeuLeuGluMet---AlaGluGluLysThrValArgAspLys-----GluLeuGluG 459
: : : : :
Db 5446 CCGGACGAGCTCGAGCGGCGAGAACAGGAGCTCCGAGCAAGCTCCACGAGATGAGGG 5505
QY 459 y-----LeuGlnValLysIleGlnArgLe 467
: : : : :
Db 5506 GGCCTCAAGTCCAAGTCCAACTCAAGTCCAGTCCAGCGGCTGGAGGCCAAGATTGCAGCT 5565
QY 467 uGluLysLeu-----CysArgAlaLeuGlnThrGluArgAsnAspLeu-----481
: : : : :
Db 5566 GGAGGAGAGCTCGACAGAGGCGCCAGAGAAACAGGCGGCCACCAAGTCGCTGAAGCA 5625
QY 482 ---AsnLysArgValGlnAspLysSerAlaGlyGlyGlnGlySerLeuThrAsp-----498
: : : : :
Db 5626 GAAAGACAGAGCTGAAGGAATCTT-CTGCAGGTGGAGGACGAGCGCAAGATGGCGG 5684
QY 499 -----SerGlyProGluArgArgProGluGlyProGlyAlaGlnAlaProSe 514
: : : : :
Db 5685 AGCAGTACAGGAGGAGGAGGAGAGAAAGGCAATGCCAGGCTCAAGCAGCTCAAGAGCAGC 5744
QY 514 rSerProArgValThrGluAlaProCysTyProGlyAlaProSerThrGluAlaSerG 534
: : : : :
Db 5745 TGAGGAGGCGAGGAGGAGTCCAGCGCATCAAGCCCAACCGCAGGAAAGCTGCAGCGG 5804
QY 534 yGln-----ThrGlyPro-----GlnG 540
: : : : :
Db 5805 AGCTGATGAGGCCACGAGAGCAACGAGGCCATGGCGCGAGGTGAACGCACTCAAGA 5864
QY 540 uProThrSerAlaArg 545
: : : : :
Db 5865 GCAAGCTCAGCGGAGG 5880

RESULT 44

US-10-116-802-14/c
; Sequence 14. Application US/10116802
; Publication No. US20030065157A1
; GENERAL INFORMATION:
; APPLICANT: Amy Lasek
; TITLE OF INVENTION: GENES EXPRESSED IN LUNG CANCER
; FILE REFERENCE: PA-0045 US
; CURRENT APPLICATION NUMBER: US/10/116,802
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/281,593
; PRIOR FILING DATE: 2001-04-04
; NUMBER OF SEQ ID NOS: 519
; SOFTWARE: PERL Program
; SEQ ID NO 14
; LENGTH: 11065
; TYPE: DNA
; ORGANISM: Homo sapiens

FEATURE:
 NAME/KEY: misc_feature
 OTHER INFORMATION: Incyte ID No: 1382907.35
 NAME/KEY: unsure
 LOCATION: 6438, 6453, 6780
 OTHER INFORMATION: a, t, c, g, or other
 US-10-116-802-14

Alignment Scores:
 Pred. No.: 3,4e-12 Length: 11065
 Score: 284.00 Matches: 150
 Percent Similarity: 37.69% Conservative: 101
 Best Local Similarity: 22.52% Mismatches: 231
 Query Match: 10.20% Indels: 185
 DB: 13 Gaps: 25

US-10-023-523-44 (1-546) x US-10-116-802-14 (1-11065)

QY 9 GlyAlaAlaLysGlnSerAsnProLysSerProGlyGlnProGluAlaGlyProGlu 28
 DB 9670 GGAGCTGCTTCAAGAAGAAACCGGCGAGAGCTCAACGCTGTCTACGAAGCTGCGCCAGCT 9611
 QY 29 GlyAlaGlnGluArgProSerGlnAlaAlaProAlaValGluAlaGluGlyProGlySer 48
 DB 9610 GGA-----GGAGGA 9602
 QY 49 SerGlnAlaProArgLysProGluGly---AlaGlnAlaArgThrAlaGlnSerGlyAla 67
 DB 9601 GGGGAACAGCTGCAAGACAGCTGGACGAGAGAGTGGAGGCGCAAGCAAGCTGGAGCG 9542
 QY 68 LeuArgAspValSerGluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTy-Cys 87
 DB 9541 CCACATCTCCACTCTCAACATCCAGCTCTCCGACTCGAAGAAGAGCTGCGAGGACTTTGC 9482
 QY 88 ValAspAsnAsnGlnGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluPro 107
 DB 9481 CAGCACCTGGAGAGCTCTGGAGAGGGAGAGAGGTTCAGAGGAGATCGAGAACCT 9422
 QY 108 -----GluAspAlaGluLys 112
 DB 9421 CACCAGCAGTACGAGGAGAGAGCGCGCGTATGATTAACATGGGAAGACCAAGAACAG 9362
 QY 113 SerArgThrTyrValAlaArgAsn-----GlyGluProGluProThrProVal 128
 DB 9361 GCTTCAGCAGAGCTGGACGCTCGTGTGTTGATTTGGACCAACACGCGCACTCGTGTGTC 9302
 QY 129 ValAsnGlyGluLysGluPro-----SerLysGlyAspProAsnThr 142
 DB 9301 CAACCTGGAAAGAGCAGAGAGGAAATTTGATCAGTTGTTAGCGGAGGAGAAACATCTC 9242
 QY 143 GluGluIleArgGlnSerAspGluValGly-AspArgAspHisArgArgProGlnGluLys 162
 DB 9241 TTCCAAATACGC-----GGATGAGAGGACAGAGCTGAGGAGAGCCAG 9197
 QY 162 stLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThrLe 182
 DB 9196 GGAGAG----- 9190
 QY 182 uSerThrProGluGluLysLeuAlaAlaLeuCysLysLysTyrAlaGluLeuLeuGluL 202
 DB 9189 -----GAACCAAGGCCCTCTCCCTGGCTCGGCGCTTGAAGAGGCTTGGAGC 9140
 QY 202 uHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValG 222
 DB 9139 CAAGAGGAGTCTGAGCGGACCAACAAATGCTCAAGCCGGAATGGAGACCTGGTCAG 9080
 QY 222 nGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuG 242
 DB 9079 CTCCAAAGATGATCGTGGGCAAGACGCTCCATGAGCTGGAGAGAGTCCAAAGCGGGCCCTGA 9020
 QY 242 uSerLeuCysArgGlyGluGlnArgHisAsnArgSerLeuLysGluGlyValGlnArg 262
 DB 9019 GACCCAGATGGAGGAGATGAGACCGCAGCTGGAGAGCTGGAGGACGAG-----CTGCA 8966

QY 262 SAlaArgGluGluGluLysArgLysGluValThr-----SerHisph 277
 DB 8965 AGCCACGGAGGAGCCAACTGGGCTGGAAGTCAACATGCAGGCGCTCAAGGGCCAGTT 8906
 QY 277 eGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLys 297
 DB 8905 CGAA-----AGGATCTCAAGCCCGGAGCAGCAGCAATGAGGAGAGAGGAGCA 8855
 QY 297 stLeuArgGlnGluAsnMetGlu----- 304
 DB 8854 ACTGCAGACAGCTTCACAGTATGAGACCGAACTGGAAGACGAGCGCAACCAAGCTGC 8795
 QY 305 -LeuAlaGluArgLeuLysLysLeuIleGlu-----GlnTyrGluLeu----- 318
 DB 8794 CTTGGCAGCTGCAGCAAGAGAGCTGGAAGGGGACCTGAAAGACCTGGAGCTTCAGCG 8735
 QY 319 -----ArgGluGluHisIleAspLysValPheLys----- 328
 DB 8734 CGACTCTGCATCAAGGGGAGGAGAGCCATCAAGCAGCTACGCAAACTCGAGGTCA 8675
 QY 329 -HisLysAspLeuGlnGlnLeuValAspAlaLys----- 340
 DB 8674 GATGAGGACTTTCAAAGAGAGCTGGAAGATGCCCGTCTCCAGAGATGAGATCTTTGC 8615
 QY 341 -----LeuGlnAlaGlu 345
 DB 8614 CACAGCCAAAGAGAATGAGAAGAAACCAAGAGCTTGAAGCAGAGACCTCATCGACTCA 8555
 QY 345 nGluMetLeuLysGluAlaGluArgHisGlnArg-----GluLysAspPheLe 362
 DB 8554 AGAGGACCTCGCGCCCTGAGAGGCTCGCAACAAAGCGGACCTCGAGAGGAGAACT 8495
 QY 362 uLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHi 382
 DB 8494 GGCAGAGGAGCTGGCCAGTAGCTGTGCGGAAGAGAACGCACTCCAGGACGAGAGCGCG 8435
 QY 382 stLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSe 402
 DB 8434 CTTGGAGGCGCGATCGCCAGCTGGAGGAGAGCTGGAGGAGGAGCGGCAACATGGA 8375
 QY 402 rLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysI 422
 DB 8374 GGCCATGAGCGAC-----CGGTCCGCAAAAGCCACACAGCAGCGC 8336
 QY 422 eLysLysLeuGluLysGluThrThrMetTyrArgSerArgTTPGluSerSerAsnLysAl 442
 DB 8335 CGAGCAGCTCAGCAACGAGCTGCCACAGCGCAGCAGCGCCAGAGAAATGAGAGTGC 8276
 QY 442 aLeuLeuGluMet---AlaGluGluLysThrValArgAspLys-----GluLeuGlu 459
 DB 8275 CCGGCGAGCTCGAGCGGCGAGCAACAGAGCTCGGAGCAAGCTCCACAGATGGAGGG 8216
 QY 459 Y-----LeuGlnValLysIleGlnArgLys 467
 DB 8215 GGCGCTCAAGTCCAAAGTTCACCATCCAGCTCGCGCTGGAGCGCAAGATTGCACAGCT 8156
 QY 467 uGluLysLeu-----CysArgAlaLeuGlnThrGluArgAsnAspLeu----- 481
 DB 8155 AGAGGAGAGCTGCAGCAGGAGCGCCAGAGAAACAGCGGCGCCACCAAGTCTCGTAGCA 8096
 QY 482 ----AsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAsp----- 498
 DB 8095 GAAAGACAAAGACTCAAGGAAATCTT-GCTGAGGTGGAGGAGCGAGCGCAAGATGGCGG 8037
 QY 499 -----SerGlyProGluArgArgProGluGlyProGlyAlaGlnAlaProSe 514
 DB 8036 AGCAGTACAGGAGCGGAGAGAAAGCAATGCCAGGCTCAGGAGCTCAGAGGCGCAGC 7977
 QY 514 rSerProArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAlaSerG 534
 DB 7976 TGGAGGAGCGCAGAGGAGGAGTCCAGCGCATCAACGCCAACCGCAGGAAGTGCAGCGGG 7917


```

QY 534 yGln-----ThrGlyPro-----GlnG 540
Db 7916 AGCTGGATGAGCCACGAGAGCAAGAGGCGGCGGCGGAGTGAACGCTCAAGA 7857

QY 540 uProThrSerAlaArg 545
Db 7856 GCAAGCTCAGCGGAGG 7841

RESULT 45
US-09-927-597-1
; Sequence 1, Application US/09927597
; Publication No. US20030032018A1
; GENERAL INFORMATION:
; APPLICANT: Malik, Fady
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; APPLICANT: Craven, Andrew
; APPLICANT: Sakowicz, Roman
; APPLICANT: Hartman, James
; TITLE OF INVENTION: Human smooth muscle myosin heavy chain
; FILE REFERENCE: CYTOPO18
; CURRENT APPLICATION NUMBER: US/09/927,597
; CURRENT FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5835
; TYPE: DNA
; ORGANISM: Human
US-09-927-597-1

Alignment Scores:
Pred. No.: 1,938-12 Length: 5835
Score: 283.00 Matches: 144
Percent Similarity: 38.22% Conservative: 101
Best Local Similarity: 22.46% Mismatches: 226
Query Match: 10.16% Indels: 171
DB: 10 Gaps: 23

US-10-023-523-44 (1-546) x US-09-927-597-1 (1-5835)
QY 9 GlyAlaAlaLysGlnSerAsnProLysSerProGlyGlnProGluAlaGlyProGlu 28
Db 3984 GGAGCTGCTTCAAGAGAAACCCGAGAGCTCAACGTGCTACGAAGCTCGCCAGCT 4043
QY 29 GlyAlaGlnGluArgProSerGlnAlaAlaProAlaValGluAlaGluGlyProGlySer 48
Db 4044 GGA-----CGAGA 4052

QY 49 SerGlnAlaProArgLysProGluGly-----AlaGlnAlaArgThrAlaGlnSerGlyAla 67
Db 4053 GCGGAACAGCTGCGAGACAGCTGACGAGGAGATGAGGCCAAGCAGAGAACTGGAGCG 4112
QY 68 LeuArgAspValSerGluGluLeuSerArgGlnLeuGluAlaAspIleLeuSerThrTyrCys 87
Db 4113 CCACATCTCCACTCTCAACATCCAGCTCTCCGACTCGAAGAGAGAGCTCGAGGACTTTGC 4172
QY 88 ValAspAsnGlnGlyGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluPro 107
Db 4173 CAGCACCCTGGAGCTCTGGAACAGCGGAGAGAGGTTCCAGAGGAGATCGAGAACT 4232
QY 108 -----GluAspAlaGluLys 112
Db 4233 CACCCAGAGTACGAGGAGAGCGCGCTTATGATAAAGTGGAAAGAGCAAGAAACAG 4292
QY 113 SerArgThrTyrValAlaArgAsn-----GlyGluProGluProVal 128
Db 4293 GCTTCAGAGAGGAGTGGAGACCTGGTGTGATTGGACACACCGGCACTCGTGTC 4352
QY 129 ValAsnGlyGluLysGluPro-----SerLysGlyAspProAsnThr 142
Db 4353 CAACCTGGAAGAGAGAGGAGAAATTTGATCAGTTGTTAGCCGAGGAGAGAAACATCTC 4412

```

```

QY 143 GluGluIleArgGlnSerAspGluValGly-AspArgAspHisArgArgProGlnGluLys 162
Db 4413 TTCCAAATAGC-----GGATGAGAGGACAGAGCTTGAGGACAGCCAG 4457

QY 162 sLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThrLe 182
Db 4458 GGAGAAG----- 4464

QY 182 uSerThrProGluLysLeuAlaAlaLeuCysLysLysTyrAlaGluLeuLeuGluL 202
Db 4465 -----GAAACCAAGCCCTCTGCTGGCTGGCCCTTGAAGAGCCCTTGAAGC 4514

QY 202 uHisArgAsnSerGlnLysGlnMetLysLeuGlnLysLysGlnSerGlnLeuValG 222
Db 4515 CAAAGAGAACTCGAGCGGACCAAAATGCTCAAAGCCGAAATGGAAGACCTGTGTG 4574

QY 222 nGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaAspSerLysLeuG 242
Db 4575 CTCRAAGGATGACGTGGCAAGAACCTCCATGAGCTGGAGAGTCCCAAGCGGCCCTGGA 4634

QY 242 uSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGluGlyValGlnAr 262
Db 4635 GACCCAGATGGAGGAGATGAAGACGAGCTGGAGAGCTGGAGGACGAG-----CTGCA 4688

QY 262 gAlaArgGluGluGluLysArgLysGluValThr-----SerHisPh 277
Db 4689 AGCCACGAGGACGCCAAACTCGCGCTGGAAGTCAACATGTCAGGCGCTCAAGGGCCAGTT 4748

QY 277 eGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLys 297
Db 4749 CGAA-----AGGATCTCCAAGCCCGGAGGAGCAATGAGGAGAGGAGGCA 4799

QY 297 sLeuArgGlnGluAsnMetGlu----- 304
Db 4800 ACTGCAGAGACACTTCACAGATGATGAGACGAACTGGAGACGAGCAAGCAAGCACTGC 4859

QY 305 -LeuAlaGluArgLeuLysLysLeuIleGlu-----GlnTyrGluLeu----- 318
Db 4860 CTTGGCAGCTGCAGCAAGAAAGAAAGTGAAGGGGACCTGAAAGACCTGGAGCTTCAGGC 4919

QY 319 -----ArgGluGluHisIleAspLysValPheLys----- 328
Db 4920 CGACTCTGCCATCAAGGGGAGGAGGAGCAATCAAGCAGACTACGCAAACTCGAGGCTCA 4979

QY 329 -HisLysAspLeuGlnGlnLeuValAspAlaLys----- 340
Db 4980 GATGAGAGACTTTCAAGAGAGCTGGAAGATGCCCTGCCCTCCAGAGATGAGATCTTGC 5039

QY 341 -----LeuGlnGlnAlaG 345
Db 5040 CACAGCCAAAGAGAAATGAGAAGAAAGCAAGAGCTTGAAGACAGACTCATGAGCTACA 5099

QY 345 nGluMetLeuLysGluAlaGluGluArgHisGlnArg-----GluLysAspPheLe 362
Db 5100 AGAGGACCTCGCCCGCTGAGAGGGCTGCCAAACAAAGCGGACCTCGAGNAGAGGAACT 5159

QY 362 uLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHi 382
Db 5160 GGCAGAGGAGCTGGCCAGTAGCTGTGCGGAAGAAACGCACTCCAGGACGAGAAGCGCCG 5219

QY 382 sLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSe 402
Db 5220 CTTGGAGCCCGGATCGCCAGCTGGAGGAGAGCTGGAGGAGGAGGAGGAGGAGGAGGAG 5279

QY 402 rLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysI 422
Db 5280 GGCCATCAGCGAC-----CGGGTCCGCAAAAGCCACAGCAGCGC 5318

QY 422 eLysLysLeuGluLysGluThrThrMetTyrArgSerArgTrpGluSerSerAsnLysAl 442
Db 5319 CGAGCAGCTCAGCAACGAGCTGGCCAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5378

QY 442 aLeuLeuGluMet--AlaGluGluLysThrValArgAspLys-----GluLeuGluL 459

```



```
QY 329 -HisLysAspLeuGlnGlnLeuValAspAlaLys----- 340
Db 5047 GATGAGGACTTTCAAAGAGAGCTGGAAGATGCCCGTCCAGAGATCAGATCTTTCG 5106
QY 341 -----LeuGlnGlnAlaGl 345
Db 5107 CACAGCCAAAGAGAATGAGAAAGCAAGAGCTTGGAAAGCAGACCTCATCGAGCTACA 5166
QY 345 nglMetLeuLysGlnAlaGluLysArgHisGlnArg-----GluLysAspPheLe 362
Db 5167 AGAGGACCTCCCGCTGAGAGGGCTCGCAAAACAGCGGACCTCGAAGAGGAGGAAC 5226
QY 362 uLeuLysGlnAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnThrHi 382
Db 5227 GCGAGAGAGCTGCCAGTACCTCTCGGAGAGAACCGCACTCCAGCAGAGAGCCCG 5286
QY 382 sLeuLysGlnGlnLeuAlaLeuThrThrGluLysPheGluPheGlnAsnThrLeuSe 402
Db 5287 CCTGGAGGCCCGGATCGCCAGCTGGAGGAGAGCTGGAGGAGGAGGAGGAGGAGGAG 5346
QY 402 rLysSerSerGluValPheThrThrPheLysGlnGlnMetGluLysMetThrLysLysI 422
Db 5347 GCGCATGAGCCAC-----CGGTCCGCAAGCCACACAGCAGGCG 5385
QY 422 eLysLysLeuGluLysGluThrThrMetTyArgSerArgTrpGluSerSerAsnLysAl 442
Db 5386 CGAGCAGCTCAACCAAGAGCTGGCCACAGAGCGCAGCAGCGCCCAAGAGATGAGAGTGC 5445
QY 442 aLeuLeuGluMet---AlaGluGluLysThrValArgAspLys-----GluLeuGlu 459
Db 5446 CCGCAGCAGCTCGAGCGCGCAGAACAGGAGCTCGGAGCAAGCTCCAGCAGATGGAGGG 5505
QY 459 Y-----LeuGlnValLysIleGlnArgLe 467
Db 5506 GCGCGTCAAGTCCAAAGTTCAAGTCCACCATCGCGCGCTGGAGGCCAGATTCACAGCT 5565
QY 467 uGluLysLeu-----CysArgAlaLeuGlnThrGluArgAsnAspLeu----- 481
Db 5566 GGAGGAGCAGGTTCGAGCAGGAGGCGCAGAGAAACAGCGCGCCACCAAGTCCGCTGAAGCA 5625
QY 482 ---AsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAsp----- 498
Db 5626 GAAAGACAAGAGCTGGAAGGAATCTT-GCTGAGGTGGAGAGCGAGCGCAAGATGGCCG 5684
QY 499 -----SerGlyProGluArgArgProGluGlyProGlyAlaGlnAlaProSe 514
Db 5685 ACCAGTACAGAGCAGCAGCAGAGAAAGCAATGCCAGGCTCAAGCAGCTCAAGAGGCGAG 5744
QY 514 rSerProArgValThrGluAlaProCysTyProGlyAlaProSerThrGluAlaSerGl 534
Db 5745 TGGAGGAGGCGAGGAGGAGTCCAGCGCATCAACGCCAACCGCAGGAAGTGCAGCGGG 5804
QY 534 Y 534
Db 5805 A 5805
RESULT 47
US-09-954-456-1602
; Sequence 1602, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
```

```
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1602
; LENGTH: 3388
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1602

Alignment Scores: 1.93e-12 Length: 3388
Pred. No.: 19 SerProGlyGlnPro-----GluAlaGlyProGluGlyAlaGlnGluArgPro 34
Score: 279.50 Matches: 146
Percent Similarity: 38.90% Conservative: 101
Best Local Similarity: 22.99% Mismatches: 225
Query Match: 10.04% Indels: 164
DB: 25 Gaps: 25

US-10-023-523-44 (1-546) x US-09-954-456-1602 (1-3388)
QY 19 SerProGlyGlnPro-----GluAlaGlyProGluGlyAlaGlnGluArgPro 34
Db 1290 GCTCCAGGACACCCAGGAGTGTCTTCAAGAGAAACCCGCGAGAGCTCAA---CGTGT 1346
QY 35 SerGlnAlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLys 54
Db 1347 TAGGAAGCTCGCGCAGCT-----GGAGGAGGAGCGGAACAGCCTCGCAAGA 1391
QY 55 ProGluGly---AlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGlu 73
Db 1392 CCAGCTGGAGGAGAGATGAGGCCAAGAGAACCTGGAGCGCCACATCTCCACTCTCAA 1451
QY 74 GluLeuSerArgGlnLeuGluAspIleLeuSerThrTyCysValAspAsnAsnGlnGly 93
Db 1452 CATCCAGCTCTCCGACTCGAAGAAAGAGCTTCGAGGACTTTGCCAGCAGCCGCGAAGCTCT 1511
QY 94 GlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluPro----- 107
Db 1512 GGAAGAGGGAAGAGAGGTTCCAGAGGAGATCGAGAACCTCACCAGCAGTACGAGGA 1571
QY 108 -----GluAspAlaGluLysSerArgThrThrTyValAla 118
Db 1572 GAAGGCGCGCTTATGATAAACTGGAAGAAAGACCAAGACAGCCTTCAGCAGGAGCTGGA 1631
QY 119 ArgAsn-----GlyGluProGluProThrProValValAsnGlyGluLysGlu 134
Db 1632 CGACCTGGTTGTGATTGGACCAACAGCGGCAACTCGTGCCAACTGGAAAGAGCA 1691
QY 135 Pro-----SerLysGlyAspProAsnThrGluGluLeuArgGlnSer 148
Db 1692 GAGGAATTTGATCAGTTGTAGCCGAGGAGAAAAACATCTCTTCCAAATACGC----- 1745
QY 149 AspGluValGly-AspArgAspHisArgArgProGlnGluLysLysAlaLysGlyLe 168
Db 1746 -----CGATGAGAGGAGCAGAGCTGAGGAGAGAGCCAGGAGGAAG----- 1785
QY 168 uGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGly 188
Db 1786 -----GAAACCAA 1793
QY 188 sLeuAlaAlaLeuCysLysLysTyAlaGluLeuLeuGluHiArgAsnSerGlnLys 208
```

Db 1794 GGCCTGTCTGCTGGCTGGCCCTTGAAGAGCGCTTGAAGCAAGAGGAACTCGAGCG 1853
 Qy 208 sGlnMetLysLeuLeuGlnLysSerGlnLeuValGlnGluLysAspHisLeuAr 228
 Db 1854 GACCAACAATGCTCAAGCGCAATGAAGAACTGTGTCAGCTCAAGATGACGTGG 1913
 Qy 228 gGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLysLeuGluLe 248
 Db 1914 CAAGAACGTCATGCTGAGTGAAGTCCCAAGCGGCGCTGAGAGCCAGATGAGAGAT 1973
 Qy 248 uGlnArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGluGluGl 268
 Db 1974 GAAGACGACGCTGGAAGAGCTGGAGACGAG-----CTGCAAGCTCGAGGAGCGCAA 2027
 Qy 268 uLysArgLysGluValThr-----SerHisPheGlnValThrLeuAsnAs 283
 Db 2028 ACTGCGGCTGGAAGTCAACATGACGCGCTCAAGGCGCAGTTTGA-----AGGA 2078
 Qy 283 pIleGlnLeuGlnMetGluGlnHisAsnGlnArgAsnSerLysLeuArgGlnGluAsnMe 303
 Db 2079 TCTCAAGCCCGGAGCGAGCAATGAGGAGAGGAGGCAACTGCGAGACAGCTTCA 2138
 Qy 303 tGlu-----LeuAlaGluArgLeuLy 310
 Db 2139 CAGATGAGACGGAAGTGAAGACGAGCGGAACGACGTCCTGGCAGCTGAGCAAA 2198
 Qy 310 sLysLeuIleGlu-----GlnTyrGluLeu 318
 Db 2199 GAAGAAGCTGGAAGGAGGACCTGAAAGACCTGAGCTTCAGGCGGACTGTCATCAAGGG 2258
 Qy 319 -ArgGluGluHisIleAspLysValPheLys-----HisLysAspLeuGlnGl 334
 Db 2259 GAGGAGGAGGAGCCATCAACAGCTACGCAACTGCGAGCTCAGTGAAGGACTTCAAAG 2318
 Qy 334 nGlnLeuValAspAlaLys----- 340
 Db 2319 AGAGCTGGAAGATGCCGCTGCTCCAGAGATGAGATCTTTCACAGCAAGAGATGA 2378
 Qy 341 -----LeuGlnAlaGlnGluMetLeuLysGluAl 351
 Db 2379 GAAGAAAGCCAAAGACTTGAAGACAGACTCATGCACTACAAGAGGACCTCGCGCGCG 2438
 Qy 351 aGluGluArgHisGlnArg-----GluLysAspPheLeuLeuLysGluAlaValGl 368
 Db 2439 TGAGAGGCTCCCAACAAAGCGGACCTCGAGAGGAGGAACTGGCAGAGAGTGGCCAG 2498
 Qy 368 uSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLeuLysGlnLeuAl 388
 Db 2499 TAGCCTGTGCGGAAGAAAGCACTCCAGAGCAAGAGCGCGCTGGAGGCGCGATCGC 2558
 Qy 388 aLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerGluValPh 408
 Db 2559 CCAGCTGAGGAGGAGCTGGAGGAGGACGCGGCACTGAGGAGCCATGACGCGAC----- 2613
 Qy 408 eThrThrPheLysGlnGluMetGluLysMetThrLysLysIleLysLysLeuGluLysGl 428
 Db 2614 -----CGGCTCCGCAAGCAAGCCACACAGCAGGCGCGGAGCTCAGCAACGA 2657
 Qy 428 uThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeuGluMet---Al 447
 Db 2658 GCTGCCACAGAGCGGCGCGCTGGAGCGCAAGAAAGAGAGTCCCGGCGGAGCTCGAGCG 2717
 Qy 447 aGluLysThrValArgAspLys-----GluLeuGluGly 459
 Db 2718 GCAGAAACAGGAGCTCCGAGCAAGCTCCACAGATGAGGAGGCGCGCTCAAGTCAAGTT 2777
 Qy 460 -----LeuGlnValLysIleGlnArgLeuLysLeu----- 470
 Db 2778 CAAGTCCACCATCGCGCGCTGGAGCGCAAGATTCACAGCTGGAGGAGGAGCTCGAGCA 2837
 Qy 471 ---CysArgAlaLeuGlnThrGluArgAsnAspLeu-----AsnLysArgValGl 486

Db 2838 GGAGGCCAGAGAGAAACAGGACGCCACCAAGTCGCTGAAGACAGAAAGCAAGAGCTGAA 2897
 Qy 486 nAspLeuSerAlaGlyGlyGlnGlnGlySerLeuThrAsp-----SerGl 500
 Db 2898 GGAAATCTT-GCTGCAGGTGGAGACGAGCGCAAGATGCCCGGAGCAGTACAGGAGCAGG 2956
 Qy 500 yProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGl 520
 Db 2957 CAGAGAAAGCAATGCGAGGTCAAGCAGCTCAAGAGCGCAGCTGGAGGAGGAGGAGGAGG 3016
 Qy 520 uAlaProCysTyrProGlyAlaProSerThrGluAlaSerGly 534
 Db 3017 AGTCCAGCGCTCAACAGCGCAACCGCAGGAAGCTGCGAGCGGA 3059

RESULT 48
 US-09-967-768A-245
 : Sequence 245, Application US/09967768A
 : Patent No. US20020150877A1
 : GENERAL INFORMATION:
 : APPLICANT: Augustus, Meena
 : TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
 : TITLE OF INVENTION: Sets
 : FILE REFERENCE: 689290-72
 : CURRENT APPLICATION NUMBER: US/09/967,768A
 : CURRENT FILING DATE: 2001-09-28
 : PRIOR APPLICATION NUMBER: US/60/236,109
 : PRIOR FILING DATE: 2000-09-28
 : PRIOR APPLICATION NUMBER: US/60/236,034
 : PRIOR FILING DATE: 2000-09-28
 : PRIOR APPLICATION NUMBER: US/60/236,111
 : PRIOR FILING DATE: 2000-09-28
 : NUMBER OF SEQ ID NOS: 325
 : SOFTWARE: PatentIn version 3.0
 : SEQ ID NO 245
 : LENGTH: 3388
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 US-09-967-768A-245

Alignment Scores:
 Pred. No.: 1,93e-12 Length: 3388
 Score: 279.50 Matches: 146
 Percent Similarity: 38.90% Conservative: 101
 Best Local Similarity: 22.99% Mismatches: 225
 Query Match: 10.04% Indels: 164
 DB: 9 Gaps: 25

US-10-023-523-44 (1-546) x US-09-967-768A-245 (1-3388)

Qy 19 SerProGlyGlnPro-----GluAlaGlyProGluGlyAlaGlnGluArgPro 34
 Db 1290 GCTCCAGGACACCCAGGAGTGTCTTCAAGAGAAACCCGCGCAGAAAGCTCAA---CGTGTC 1346
 Qy 35 SerGlnAlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLys 54
 Db 1347 TAGCAAGCTGCGCCAGCT-----GGAGGAGGAGCGGAACAGCCTCAAGA 1391
 Qy 55 ProGluGly---AlaGlnAlaArgThrAlaGlnSerGlyValaLeuArgAspValSerGlu 73
 Db 1392 CAGCTGGAGGAGATGGAGCCAGCAGAACCTGGAGCGCCACATCTCCACTCTCAA 1451
 Qy 74 GluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysValaAspAsnGlnGly 93
 Db 1452 CATCCAGCTCTCCGACTCGAAGAGAGCTGCGAGGACTTTGCCAGCACCGTGAAGCTCT 1511
 Qy 94 GlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluPro----- 107
 Db 1512 GGAGAGGGGAGAGAGAGGTTCCAGAGGAGATCCAGAACCTCACCAGCAGTACGAGGA 1571
 Qy 108 -----GluAspAlaGluLysSerArgThrTyrVala 118
 Db 1572 GAAGCGCGCGCTTATGATAAACTGGAAGAACCAAGAGGCTTTCAGCAGGAGCTGGA 1631

QY 119 ArgAsn-----GlyGluProGluProThrProValValAsnGlyGluLysGlu 134
Db |||||
1632 CGACCTGTTGTTGATTTGGACACACCGCGCAACTCGTGTCCAACTGGAAAGAGCA 1691
QY |||||
135 Pro-----SerLysGlyAspProAsnThrGluGluLeuArgGlnSer 148
Db |||||
1692 GAGGAATTTGATCAGTTGTTAGCGGAGGAGAAACATCTCTTCCAAATACGC----- 1745
QY |||||
149 AspGluValGly-AspArgAspHisArgArgProGlnGluLysLysLysLysGlyLe 168
Db |||||
1746 -----GGATGAGAGGGGACAGAGCTGAGCGCAGAGCCAGCGGAGAG----- 1785
QY |||||
168 uGlyLysGluLeuThrLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLys 188
Db |||||
1786 -----GAAACCAA 1793
QY |||||
188 sLeuAlaAlaLeuCysLysLysTyrAlaGluLeuLeuGluHisArgAsnSerGlnLys 208
Db |||||
1794 GGCCTGTCTCCCTGCGGCGCTTGAAGAGAGCCCTTGAAGACCAAGAGGAACTCGAGCG 1853
QY |||||
208 sGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuAr 228
Db |||||
1854 GACCAACAAATGCTCAAGCGGAAATGGAAGACCTGTCTAGCTCCAAAGGATGACGTGGG 1913
QY |||||
228 gGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLysCysArgGluLe 248
Db |||||
1914 CAAGAACCTCCATGAGCTGGAGAGTCCCAAGCGGCGCTGGAGACCCACAGATGGAGGAT 1973
QY |||||
248 uGlnArgHisAsnArgSerLeuLysGluGlyValGlnArgAlaArgGluGluGlu 268
Db |||||
1974 GAAGACGAGCTGGAGAGCTGGAGGAGAG-----CTGCAAGCTCGGAGGAGCCAA 2027
QY |||||
268 uLysArgLysGluValThr-----SerHisPheGlnValThrLeuAsnAs 283
Db |||||
2028 ACTGCGGTGGAGTCAACATGACGCGCTCAAGGCGCAGTTCGAA-----AGGA 2078
QY |||||
283 pIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMe 303
Db |||||
2079 TCTCAAGCCCGGACGAGCAGATGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2138
QY |||||
303 tGlu-----LeuAlaGluArgLeuLys 310
Db |||||
2139 CGAGTATGAGCGGAACTGGAGAGCAGCGGAAACAACTGCTGCGGAGCTGCAGCAAA 2198
QY |||||
310 sLysLeuLeuGlu-----GlnTyrGluLeu----- 318
Db |||||
2199 GAAGAACTGGAAGGAGGAGCTGAAAGACCTGAGGCTTCAGGCGGAGCTCTGCCATCAAGG 2258
QY |||||
319 -ArgGlnGluHisLysAspLysValPheLys-----HisLysAspLeuGlnGlu 334
Db |||||
2259 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2318
QY |||||
334 nGlnLeuValAspAlaLys----- 340
Db |||||
2319 AGAGCTGGAAGATGCGCGTCCAGAGATGAGATCTTTGCCACAGCAAGAGAAATGA 2378
QY |||||
341 -----LeuGlnGlnAlaGlnGluMetLeuLysGluAl 351
Db |||||
2379 GAAGAAAGCAAGAGCTTGAAGCAGACCTATGAGCTTACAGAGGAGGAGGAGGAGGAGGAG 2438
QY |||||
351 aGluGluArgHisGlnArg-----GluLysAspPheLeuLysGluAlaValGlu 368
Db |||||
2439 TGAGAGGCTGCCAAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2498
QY |||||
368 uSerGlnArgMetCysGluLeuMetLysGlnGlnThrHisLeuLysGlnGlnLeuAl 388
Db |||||
2499 TAGCCTGTCCGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2558
QY |||||
388 aLeuTyrThrGluLysPheGlnGluPheGlnAsnThrLeuSerLysSerSerGluValPh 408
Db |||||
2559 CCAGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2613
QY |||||
408 eThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLysLysLysLysLys 428

Db 2614 -----CGGTCGCAAAAGCCACACAGAGGCGGAGCAGCTCAGCAACGA 2657
QY |||||
428 uThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeuGluMet---Al 447
Db |||||
2658 GTGGCCACAGAGCGGAGCAGCGCCAGAAAGATGAGAGTCCCGGACAGCTCGAGCG 2717
QY |||||
447 aGluGluLysThrValArgAspLys-----GluLeuGluGly----- 459
Db |||||
2718 GCAGAACAGGAGCTCCGAGCAAGCTCCACAGATGGAGGGGGCGGTCAAGTCCAAGTT 2777
QY |||||
460 -----LeuGlnValLysLysLysGlnArgLeuGluLysLeu----- 470
Db |||||
2778 CAAGTCCACATCGCGCGCTGGAGCCCAAGATTCCAGCTGGAGGAGCAGCGTCCAGCA 2837
QY |||||
471 ---CysArgAlaLeuGlnThrGluArgAsnAspLeu-----AsnLysArgValGlu 486
Db |||||
2838 GGAGGCCACAGAGAAACAGCGGAGCCACCAAGTCCGTGAAGCAGAGAAAGACAGAGCTGAA 2897
QY |||||
486 nAspLeuSerAlaGlyGlyGlnGlnSerLeuThrAsp-----SerGlu 500
Db |||||
2898 GGAATCTT-GCTGCAAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2956
QY |||||
500 yProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGlu 520
Db |||||
2957 CAGAAAGGCAATGCGAGGCTCAACAGCTCAAGAGGAGCTGAGGAGGAGGAGGAGGAGGAGG 3016
QY |||||
520 uAlaProCysTyrProGlyAlaProSerThrGluAlaSerGly 534
Db |||||
3017 AGTCCAGCGCATCAACGCCAACCGCAGGAAGCTGCAGCGGGA 3059
RESULT 49
US-09-954-531-988
; Sequence 988, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
; APPLICANT: Weaver, Zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cance
; FILE REFERENCE: 689290-77
; CURRENT APPLICATION NUMBER: US/09/954,531
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/60/233,133
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 1392
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 988
; LENGTH: 3388
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-531-988
Alignment Scores:
Pred. No.: 1,93e-12 Length: 3388
Score: 279.50 Matches: 146
Percent Similarity: 38.90% Conservative: 101
Best Local Similarity: 22.99% Mismatches: 225
Query Match: 10.04% Indels: 164
DB: 9 Gaps: 25
US-10-023-523-44 (1-546) x US-09-954-531-988 (1-3388)
QY 19 SerProGlyGlnPro-----GluAlaGlyProGluGlyAlaGlnGluArgPro 34
Db |||||
1290 GCTCCAGACACCCCGAGAGTGTCTTCAAGAGAAACCCGCGCAGAGCTCAA---CGTGC 1346

Wed Jun 9 08:59:05 2004

us-10-023-523-44.rnpb

Page 50

Search completed: June 8, 2004, 13:16:19
Job time : 1131.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: June 8, 2004, 06:37:59 ; Search time 943.755 Seconds
(without alignments)
2561.956 Million cell updates/sec

Title: US-10-023-523-8
Perfect score: 2702
Sequence: 1 KSPGQPEAGPEGAQRPSPQ.....APSTEASGGTGPQPTSPARA 530

Scoring table:
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Dgapop 6.0, Dext 7.0

Searched: 2995936 seqs, 2280998010 residues

Total number of hits satisfying chosen parameters: 5991872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Command line parameters:
-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10023523/runat_08062004_063750_29986/app_query.fasta_1.1422
-DB=Published Applications NA -QFMI=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=50 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=50 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HSPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10023523 @CEN 1.1 810 @runat_08062004_063750_29986
-NCPU=6 -ICPU=3 -NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US05_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09D_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09E_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

ALIGNMENTS

RESULT 1
US-09-962-055-17
; Sequence 17, Application US/09962055
; Patent No. US2002052033A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; Lees, Robert S.
; Law, Simon W.
; Arjona, Anibal A.

TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND TREATING ATHEROSCLEROSIS

NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA

	1	2702	100.0	4697	9	US-09-962-055-17	Sequence 17, Appl
	2	2702	100.0	4597	9 <td>US-09-976-740-17</td> <td>Sequence 17, Appl</td>	US-09-976-740-17	Sequence 17, Appl
	3	2702	100.0	4897	13	US-10-671-242-17	Sequence 17, Appl
	4	2702	100.0	4697	13	US-10-023-523-17	Sequence 17, Appl
	5	2702	100.0	4697	14	US-10-023-523-17	Sequence 17, Appl
	6	2702	100.0	4697	16	US-10-616-187-17	Sequence 17, Appl
	7	2693	99.7	1638	9 <td>US-09-976-740-46</td> <td>Sequence 46, Appl</td>	US-09-976-740-46	Sequence 46, Appl
	8	2693	99.7	1638	13	US-10-671-242-46	Sequence 46, Appl
	9	2693	99.7	1638	14	US-10-023-523-46	Sequence 46, Appl
	10	2693	99.7	1638	14	US-10-023-523-46	Sequence 46, Appl
	11	2693	99.7	1638	16	US-10-616-187-46	Sequence 46, Appl
	12	2693	99.7	2523	13	US-10-276-774-784	Sequence 784, App
	13	2465.5	91.2	4722	9 <td>US-09-962-055-14</td> <td>Sequence 14, Appl</td>	US-09-962-055-14	Sequence 14, Appl
	14	2465.5	91.2	4722	9 <td>US-09-976-740-14</td> <td>Sequence 14, Appl</td>	US-09-976-740-14	Sequence 14, Appl
	15	2465.5	91.2	4722	13	US-10-671-242-14	Sequence 14, Appl
	16	2465.5	91.2	4722	14	US-10-023-523-14	Sequence 14, Appl
	17	2465.5	91.2	4722	14	US-10-023-523-14	Sequence 14, Appl
	18	2465.5	91.2	4722	16	US-10-616-187-14	Sequence 14, Appl
	19	1290.5	47.8	2356	16	US-10-115-831-34	Sequence 34, Appl
	20	1285	47.6	2045	9 <td>US-09-960-253-172</td> <td>Sequence 172, App</td>	US-09-960-253-172	Sequence 172, App
	21	820.5	30.4	2194	16 <td>US-10-108-260A-1586</td> <td>Sequence 1586, Ap</td>	US-10-108-260A-1586	Sequence 1586, Ap
	22	721.5	26.7	22255	9 <td>US-09-976-740-51</td> <td>Sequence 51, Appl</td>	US-09-976-740-51	Sequence 51, Appl
	23	721.5	26.7	22255	13	US-10-671-242-51	Sequence 51, Appl
	24	721.5	26.7	22255	14	US-10-023-523-51	Sequence 51, Appl
	25	721.5	26.7	22255	14	US-10-023-523-51	Sequence 51, Appl
	26	721.5	26.7	22255	16	US-10-616-187-51	Sequence 51, Appl
	27	713	26.4	529	14 <td>US-10-040-739-869</td> <td>Sequence 869, App</td>	US-10-040-739-869	Sequence 869, App
	28	652.5	24.1	614	9 <td>US-09-833-381-1333</td> <td>Sequence 1333, Ap</td>	US-09-833-381-1333	Sequence 1333, Ap
	29	501	18.5	1955	13 <td>US-10-425-114-32431</td> <td>Sequence 32431, A</td>	US-10-425-114-32431	Sequence 32431, A
	30	478	17.7	937	12 <td>US-09-878-143-1454</td> <td>Sequence 1454, Ap</td>	US-09-878-143-1454	Sequence 1454, Ap
	31	464	17.2	334	10 <td>US-09-918-985-18899</td> <td>Sequence 18899, A</td>	US-09-918-985-18899	Sequence 18899, A
	32	454.5	16.8	1696	13 <td>US-10-424-599-64592</td> <td>Sequence 64592, A</td>	US-10-424-599-64592	Sequence 64592, A
	33	444.5	16.5	1351	13 <td>US-10-425-114-6003</td> <td>Sequence 6003, Ap</td>	US-10-425-114-6003	Sequence 6003, Ap
	34	350	13.0	405	9 <td>US-09-764-869-467</td> <td>Sequence 467, App</td>	US-09-764-869-467	Sequence 467, App
	35	350	13.0	405	15 <td>US-10-091-504-467</td> <td>Sequence 467, App</td>	US-10-091-504-467	Sequence 467, App
	36	350	13.0	405	16 <td>US-10-227-577-467</td> <td>Sequence 467, App</td>	US-10-227-577-467	Sequence 467, App
	37	318	11.8	725	16 <td>US-10-264-049-1238</td> <td>Sequence 1238, Ap</td>	US-10-264-049-1238	Sequence 1238, Ap
	38	313	11.6	981	12 <td>US-09-876-143-1445</td> <td>Sequence 1445, Ap</td>	US-09-876-143-1445	Sequence 1445, Ap
	39	304.5	11.3	1035	13 <td>US-10-423-114-18551</td> <td>Sequence 18551, A</td>	US-10-423-114-18551	Sequence 18551, A
	40	282.5	10.5	5937	10 <td>US-09-927-597-3</td> <td>Sequence 3, Appli</td>	US-09-927-597-3	Sequence 3, Appli
	41	282.5	10.5	6861	15 <td>US-10-171-311-161</td> <td>Sequence 161, App</td>	US-10-171-311-161	Sequence 161, App
	42	282.5	10.5	6861	16 <td>US-10-341-434-102</td> <td>Sequence 102, App</td>	US-10-341-434-102	Sequence 102, App
	43	282.5	10.5	11065	13 <td>US-10-116-803-14</td> <td>Sequence 14, Appl</td>	US-10-116-803-14	Sequence 14, Appl
	44	281.5	10.4	5835	10 <td>US-09-927-597-1</td> <td>Sequence 1, Appli</td>	US-09-927-597-1	Sequence 1, Appli
	45	281.5	10.4	6900	15 <td>US-10-171-311-163</td> <td>Sequence 163, App</td>	US-10-171-311-163	Sequence 163, App
	46	278.5	10.3	3388	9 <td>US-09-954-456-1602</td> <td>Sequence 1602, Ap</td>	US-09-954-456-1602	Sequence 1602, Ap
	47	278.5	10.3	3388	9 <td>US-09-967-768A-245</td> <td>Sequence 245, App</td>	US-09-967-768A-245	Sequence 245, App
	48	278.5	10.3	3388	9 <td>US-09-954-531-988</td> <td>Sequence 988, App</td>	US-09-954-531-988	Sequence 988, App
	49	278.5	10.3	3388	9 <td>US-09-954-531-1382</td> <td>Sequence 1382, Ap</td>	US-09-954-531-1382	Sequence 1382, Ap
	50	278.5	10.3	3388	10 <td>US-09-873-367C-85</td> <td>Sequence 85, Appl</td>	US-09-873-367C-85	Sequence 85, Appl

COUNTRY: USA
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/962,055
 FILING DATE: 24-Sep-2001
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/979,608
 FILING DATE: 26-Nov-1997
 APPLICATION NUMBER: US 60/031,930
 FILING DATE: 27-Nov-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Myers, Louis
 REGISTRATION NUMBER: 35,965
 REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617/542-5070
 TELEFAX: 617/542-8906
 INFORMATION FOR SEQ ID NO: 17:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 4697 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 FEATURE:
 NAME/KEY: Coding Sequence
 LOCATION: 3..1592
 SEQUENCE DESCRIPTION: SEQ ID NO: 17:

US-09-962-055-17

Alignment Scores:

Pred. No.: 6,896-210 Length: 4697
 Score: 2702.00 Matches: 530
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-10-023-523-8 (1-530) x US-09-962-055-17 (1-4697)

QY 1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluAtrProSerCln 20
 DB 3 AAACAGCCCAAGGACACCCGAGCAGGACCCGAGGAGCCAGGAGCCAGCCAG 62
 QY 21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40
 DB 63 GCGGCTCTGAGTAGAAGCAGAGAGTCCCGGCGAGCAGCCAGGCTCTCGAAGCCGGAG 122
 QY 41 GlyAlaGlnAlaArgThrAlaGlnSerClyAlaLeuArgAspValSerGluGluLeuSer 60
 DB 123 GCGGCTCAAGCCAGAACCGGCTCAGTCTGGGGCCCTTCGTGTGTCTGTGAGAGCTGAGC 182
 QY 61 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGly 80
 DB 183 CGCCAACTGGAAGACATACTGAGCACATACTGTGTGGACATAACAGGGGGGCCCCGGC 242
 QY 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
 DB 243 GAGGATGGGACACAGGGTGAGCCGCGCTGAACCCGAAGATGACAGAAAGTCCCGAACCTAT 302
 QY 101 ValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGlyLysGluProSer 120
 DB 303 GTGGCAAGGANTGGGACCTGAACCACTCCAGTAGTCTATGGAGAGAGAACCTCC 362
 QY 121 LysGlyAspProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAspHis 140
 DB 363 AAGGGGGATCCAAACACAGAGAGATCCGGCAGAGTGCAGAGGTCGAGAGCCAGACCAT 422
 QY 141 ArgArgProGlnGluLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMet 160

DB 423 CGAAGGCCACAGAGAGAGAAAAAGCCAGGGTTTGGGAGGAGATCACGTTGCTGATG 482
 QY 161 GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysTyr 180
 DB 483 CAGACATTGAATATCTCTGAGTACCCAGAGAGAGAGCTGCTCTCTGTGCAAGAAATAT 542
 QY 181 AlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLys 200
 DB 543 GCTGAATGCTGAGGAGCACCCGGAATTACAGAGAGCATGAGGCTCTCTACAAAAAAG 602
 QY 201 GlnSerGlnLeuValGlnGlnLysAspHisLeuArgGlyGluHisSerLysAlaValLeu 220
 DB 603 CAGAGCCAGCTGCTGCAAGAGAGAGACCACTCGCGGCTGAGCACAGCAAGCGCTCTG 562
 QY 221 AlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLys 240
 DB 663 CCGCGCAGCAAGCTTGAGAGCTTATCGCTGAGCTGAGCGGACCAACCGCTCCCTCAAG 722
 QY 241 GluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHis 260
 DB 723 GAAGAAGGTGTGAGCGGCGCCGAGGAGGAGAGAGCAAGCGAAGAGGTGACCTCGCAC 782
 QY 261 PheGluValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSer 280
 DB 783 TTCAGGTGACACTGAATGACATTCAGCTGAGTGGAGATGGAACAGCACAAATGAGCGCACTCC 842
 QY 281 LysLeuArgGlnGlnLeuAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTyr 300
 DB 843 AAGCTGCGCCACAGAGAACATGGAGCTGGCTGAGAGGCTCAAGAGCTGATTGAGCAGTAT 902
 QY 301 GluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeu 320
 DB 903 GAGCTCGCGAGGAGCATATCGCAAGTCTTCAAAACACAGAGACCTCAACAGCAGCTG 962
 QY 321 ValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGln 340
 DB 963 GTGGATGCCAAGCTCCAGCAGGCCAGGAGATGCTTAAGAGAGCGAGAGAGCGGCACGAG 1022
 QY 341 ArgGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMet 360
 DB 1023 CCGGAGAGAGGATTTCTCTCTGAAAGAGGAGTAGAGTCCAGAGGATGTGTGAGCTGATG 1082
 QY 361 LysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGlu 380
 DB 1083 AAGCACAAGACACCCACCTGAGCAACAGCTTGCCCTATACACAGAGAGTTTGAGGAG 1142
 QY 381 PheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGlu 400
 DB 1143 TTCAGAACACACTTTTCAAAAGCAGCAGGATTTTACCACATTCAAGCAGCAGATGGAA 1202
 QY 401 LysMetThrLysLysLysLysLysLysGluLysGluThrThrMetTyrArgSerArgTrp 420
 DB 1203 AGATGACTAAGAAAGATCAAGAGGCTGGAGAAAGAAACACCATGTACCGGTCCCGGTGG 1262
 QY 421 GluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGlu 440
 DB 1263 GAGAGCAGCAACAAGCCCTGCTTGAAGTGGCTGAGGAGAAACAGTCCGGGATAAGAA 1322
 QY 441 LeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThr 460
 DB 1323 CTGGAGGGGCTTCAGGTAAAAATCCAAAGCTGGAGAGAGCTGTGCGGGGCACTGCAGACA 1382
 QY 461 GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeu 480
 DB 1383 GAGCGCAATGACTGAACAGAGGGGTACAGGACCTGAGTGTGTGGCGCAGGGCTCCCTC 1442
 QY 481 ThrAspSerGlyProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerPro 500
 DB 1443 ACTGACAGTGGCCCTGAGAGGAGGCGCAGAGGGGCTTGGGGCTCAAGCACCCAGCTCC 1502
 QY 501 ArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThr 520

Db 1503 AGGTCACAGAGCCCTTGCTACCCAGGAGCACCGAGCAGCAGAGCATCAGGCAGACT 1562
QY 521 GlyProGlnGluProThrSerAlaArgAla 530
Db 1563 GGGCCTCAAGAGCCACCTCCGCCAGGGCC 1592

RESULT 2

US-09-976-740-17
; Sequence 17, Application US/09976740
; Publication No. US20020194633A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/976,740
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 4697
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(1592)
US-09-976-740-17

Alignment Scores:

Pred. No.: 6,89e-210 Length: 4697
Score: 2702.00 Matches: 530
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-023-523-8 (1-530) x US-09-976-740-17 (1-4697)

QY 1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 20
Db 3 AAAAGCAGCCAGGACCAACCGGAAGCAGGACCCGAGGAGCCAGGAGCCGAGCCAGCCAG 62
QY 21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40
Db 63 GCGGCTCTCGAGTAGAAGCAGAGGTCCTCCGAGCAGCCAGGCTCTCCGAGGAGCCGAG 122
QY 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 60
Db 123 GGGGCTCAAGCAGACAGCGCTCAGTCTGGGGCCCTTCGTGATGTCTCTGAGGAGCTGAGC 182
QY 61 ArgGlnLeuGluAspIleuSerThrTyrcysValAspAsnGlnGlyGlyProGly 80
Db 183 CCGCAACTCGAAGACATCTGTAGACATACCTGTGTGACAAATACACAGGGGGCCCCGGC 242
QY 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
Db 243 GAGGATGGGGCAGCAGGCTGAGCGGCTGAACCGAGAGTGCAGAGAGTCCCGGACCTAT 302
QY 101 ValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGluLysGluProSer 120
Db 303 GTGGCAAGGAATGGGAGGCTGAACCAACTCCAGTAGTCTATGGAGAGAAGAACCTCC 362

QY 121 LysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHis 140
Db 363 AAGGGGATCCAAACACAGAGAGATCCGCGACAGTACAGAGTCCGAGACCGAGCAT 422
QY 141 ArgArgProGlnGluLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMet 160
Db 423 CGAAGGCCACAGGAGAGAAAAAAGCAAGGTTTGGGGAAGGAGATCACTTGCTGATG 482
QY 161 GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysTyr 180
Db 483 CAGACATTGAATCTCTGAGTACCCAGAGGAGAGCTGCTCTCTCTGTCAGAGAGTAT 542
QY 181 AlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLys 200
Db 543 GCTGAACCTGCTGGAGGAGCACCGGAATTCACAGACAGATGAAGCTCTCTACAGAAAAA 602
QY 201 GlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeu 220
Db 603 CAGAGCCAGCTGGTGCAGAGAGAGGACCACTCCGCGGTGAGCAGCAGCAAGGCGCTG 662
QY 221 AlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLys 240
Db 663 GCCCGCAGCAGCTTGAGACCTATGCCGTGAGCTGCAGCGGCACAAACGCTCCCTCAAG 722
QY 241 GluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHis 260
Db 723 GAAGAAGTGTGCAGCGGCGCGGAGGAGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 782
QY 261 PheGlnValThrLeuAsnAspIleGlnLeuGlnMetGlnGlnHisAsnGluArgAsnSer 280
Db 783 TTCAGGTGACACTGAATGACATTCAGCTGCAGATGGAACAGCACAATGAGCGCACTCC 842
QY 281 LysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLysLysLysLys 300
Db 843 AAGCTGCGCCAGAGAACATGGAGCTGGCTGAGAGGCTCAAGAAGCTGATTGAGCAGTAT 902
QY 301 GluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnGlnLeu 320
Db 903 GAGCTGCGCGAGGAGCATATCGACAAAGTCTTCAACACAGAGGAGCCTTACACACAGCAG 962
QY 321 ValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluGluArgHisGln 340
Db 963 GTGGATGCCAAGCTCCAGAGGCCAGGAGATGCTAAAGAGGAGGAGGAGGAGGAGGAG 1022
QY 341 ArgGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMet 360
Db 1023 CCGGAGAAGGATTTCTCTCTGAAAGAGGAGGAGTGTAGTCCAGAGGAGTGTGTGAGCTG 1082
QY 361 LysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGlu 380
Db 1083 AAGCAGCAAGAGAGACCCACCTGAAAGCAACAGCTTCCCTCTATACAGAGAGAGTTTGA 1142
QY 381 PheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGlu 400
Db 1143 TTCCAGAACACACTTTCACAAAGCAGCGAGGTATTCCACCACTTCAAGCAGGAGAGTGA 1202
QY 401 LysMetThrLysLysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArgTyr 420
Db 1203 AAGATGACTAAGAGAGATCAAGAGCTCGAGAGAGAAACCACTGATACCCGGTCCCGGTG 1262
QY 421 GluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGlu 440
Db 1263 GAGAGCAGCAACAGAGGCTTCTTGTAGATGGCTGAGAGAGAAACAGTCCGGGATAAGAA 1322
QY 441 LeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThr 460
Db 1323 CTGGAGGCGCTGCGAGGTAAAAATCCAAACGGCTGGAGAGAGTGTGCCGGGACCTGCAACA 1382
QY 461 GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeu 480
Db 1383 GAGCGCAATGACCTGAACAGAGAGGTACAGGACCTGAGTGTGTGTGCGCCAGGCGCTCC 1442

QY 481 ThrAspSerGlyProGluArgArgProGluGlyProGluGlyValAlaGlnAlaProSerSerPro 500
 Db 1443 ACTGACAGTGGCCCTGAGAGAGAGCCAGAGGGGCTGGGGCTCAAGACCCAGCTCCCC 1502
 QY 501 ArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThr 520
 Db 1503 AGGCTCACAGAGCGCTTCTCTCCAGGAGCACCAGAGCAGCAGAGCATCAGCCAGACT 1562
 QY 521 GlyProGlnGluProThrSerAlaArgAla 530
 Db 1563 GGGCTCAAGAGCCACCTCCGAGGCCC 1592

RESULT 3

US-10-671-242-17
 ; Sequence 17, Application US/10671242
 ; Publication No. US20040040049A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lees, Ann M.
 ; APPLICANT: Lees, Robert S.
 ; APPLICANT: Law, Simon W.
 ; APPLICANT: Arjona, Anibal A.
 ; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
 ; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
 ; TITLE OF INVENTION: ATHEROSCLEROSIS
 ; FILE REFERENCE: 10797-004001
 ; CURRENT APPLICATION NUMBER: US/10/671,242
 ; CURRENT FILING DATE: 2003-09-24
 ; PRIOR APPLICATION NUMBER: US/09/616,289
 ; PRIOR FILING DATE: 2000-07-14
 ; PRIOR APPLICATION NUMBER: US 09/517,849
 ; PRIOR FILING DATE: 2000-03-02
 ; PRIOR APPLICATION NUMBER: US 08/979,608
 ; PRIOR FILING DATE: 1997-11-26
 ; PRIOR APPLICATION NUMBER: US 60/031,930
 ; PRIOR FILING DATE: 1996-11-27
 ; PRIOR APPLICATION NUMBER: US 60/048,547
 ; PRIOR FILING DATE: 1997-06-03
 ; NUMBER OF SEQ ID NOS: 53
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 17
 ; LENGTH: 4697
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (3)...(1592)
 US-10-671-242-17

Alignment Scores:
 Pred. No.: 6,89e-210 Length: 4697
 Score: 2702.00 Matches: 530
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 13 Gaps: 0

US-10-023-523-8 (1-530) x US-10-671-242-17 (1-4697)

QY 1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyValAlaGlnAlaProSerGln 20
 Db 3 AAAAGAGCCCGAGCAACCGAAGCAGGAGCCAGAGGAGCCAGAGCGCCAGCCAG 62
 QY 21 AlaAlaProAlaValGluAlaGlyProGlySerSerGlnAlaProArgLysProGlu 40
 Db 63 GCGGCTCTCGAGTAGAGCAGAGGTCCCGGAGCAGCCAGGCTCTCTCGAAGCCGAG 122
 QY 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluLeuSer 60
 Db 123 GGGGCTCAAGCCAGAACCGCTAGTCTGGGGCCCTTCGTGTGTCTCTGAGGAGCTGAGC 182
 QY 61 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnGlnGlyGlyProGly 80
 Db 183 CGCCAACTGGAGACATACTGAGCACAATCTGTGTGGACATAAACAGGGGGGCCCGCGC 242

QY 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
 Db 243 CAGATGGGGACACAGGGTGAGCGGCTGAACCCGAGATGCAGAGAGATCCGGACCTAT 302
 QY 101 ValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGluLysGluProSer 120
 Db 303 GTGGCAAGGAATGGGAGCCTGAACCAACTCCAGTAGTCTATGGAGAGAGAACCCCTCC 362
 QY 121 LysGlyAspProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAsnHis 140
 Db 363 AAGGGGATCCAAACACAGAGAGATCCGGCAGAGTACCGAGGTCCGAGAGCCAGACAT 422
 QY 141 ArgArgProGlnGluLysLysAlaLysGlyLeuGlyLysGluLysGluLysLeuMet 160
 Db 423 CGAAGGCCACAGGATG 482
 QY 161 GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuLysLysTyr 180
 Db 483 CAGACATTTGAATACTCTGAGTACCCAGAGAGAGAGTGGCTGCTCTGTCAAGAGAT 542
 QY 181 AlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuGlnLysLys 200
 Db 543 GCTGAATCTGGAGGAGCACCAGGATTCACAGAGCAGATGAAGCTCCTACAGAGAG 602
 QY 201 GlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeu 220
 Db 603 CAGAGCCAGCTGGTGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 662
 QY 221 AlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLys 240
 Db 663 GCCCGCAGAGCTTTGAGAGCCTATGCGGTGAGCTGCGGCGGACACACCCCTCCTCAAG 722
 QY 241 GluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHis 260
 Db 723 GAAGAAGTGTGCAGCGGCGGCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 782
 QY 261 PheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisGlnArgAsnSer 280
 Db 783 TTCAGGTGACATGAATGACATTCAGCTGAGATGAGGAGAGAGAGAGAGAGAGAGAG 842
 QY 281 LysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuGluGlnTyr 300
 Db 843 AAGCTGCCCAAGAGACATGAGCTGGCTGAGAGGCTCAAGAAGCTGATTGAGAGAT 902
 QY 301 GluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeu 320
 Db 903 GAGCTGCGCGAGGAGCATATCGACAAAGTCTTCAACACAGAGAGAGAGAGAGAGAG 962
 QY 321 ValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGln 340
 Db 963 GTGGATGCCAAGCTCCAGCAGGCGCCAGGAGATGCTAAAGAGGAGAGAGAGAGAGAG 1022
 QY 341 ArgGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMet 360
 Db 1023 CCGGAGAAGGATTTTCTCTCAAAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1082
 QY 361 LysGlnGlnGluThrHisLeuLysGlnLeuAlaLeuTyrThrGluLysPheGluGlu 380
 Db 1083 AAGCAGCAG 1142
 QY 381 PheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGlu 400
 Db 1143 TTCAGAACACACTTTCACAAAGCAGCAGGATTCACACATTCACAGCAGAGAGAGAGAG 1202
 QY 401 LysMetThrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 420
 Db 1203 AAGATGACTAAGAGATCAAGAGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1262
 QY 421 GluSerSerAsnLysAlaLeuLeuGluMetAlaGluLysThrValArgAspLysGlu 440
 Db 1263 GAGAGCAGCAACAGGCGCTGCTTGTGATGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1322


```
Qy 361 LysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGlu 380
Db 1083 AAGCAGCAAGAGAGCCACCTGAACACAGCTTGCCTATACACAGAGAGTTTGAGGAG 1142
Qy 381 PheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGlu 400
Db 1143 TTCAGAACACACTTCCAAAGCAGCGAGGTATTCCACCATTCAAGCAGGAGATGGAA 1202
Qy 401 LysMetThrLysLysIleLeuLysLeuGluLysGluThrThrMetTyrArgSerArgTip 420
Db 1203 AAGATGACTAGAGAGATCAGAGACTGGAGAAAGAACCCACCATGTACCGGTCCCGGTGG 1262
Qy 421 GluSerSerAsnLysAlaLeuLeuGluMetAlaGluLysThrValArgAspLysGlu 440
Db 1263 CAGAGCAGCAACAAGGCCCTGCTTGCAGATGGCTGAGGAGAAAACAGTCCGGGATAAAGAA 1322
Qy 441 LeuGluGlyLeuGlnValLysIleGlnArgGlnGluLysLeuCysArgAlaLeuGlnThr 460
Db 1323 CTGAGGGCCCTGCAGGTAAAATCAACGGCTGAGAGAGCTGTGCCGGGCACCTCAGACA 1382
Qy 461 GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeu 480
Db 1383 CAGCGCAATGACCTGAACAAGAGGTACAGACCTGAGTGTGTGGCCAGGGCTCCCTC 1442
Qy 481 ThrAspSerGlyProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerPro 500
Db 1443 ACTGACAGTGGCCCTGAGAGAGGCCAGAGGGGCTGCGGCTCAAGCACCAGCTCCCCC 1502
Qy 501 ArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThr 520
Db 1503 AGGTCACAGAGAGCGCTTGTATCCAGGAGCACCGAGACCGAGCATCAGGCCAGACT 1562
Qy 521 GlyProGlnGluProThrSerAlaArgAla 530
Db 1563 GGGCTCAAGAGCCACCTCCGCCAGGGCC 1592
```

RESULT 6

```
US-10-616-187-17
; Sequence 17, Application US/10616187
; Publication No. US20040013668A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/616,187
; PRIOR FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 4697
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(1592)
US-10-616-187-17
```

Alignment Scores:

```
Pred. No.: 6,89e-210 Length: 4697
Score: 2702.00 Matches: 530
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 16 Gaps: 0
US-10-023-523-8 (1-530) x US-10-616-187-17 (1-4697)
Qy 1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyValaGlnGluArgProSerGln 20
Db 3 AAAGCAGCCAGGAGACACCGGAAGCAGGAGCCCGAGGAGCCAGGAGCGGCCAGCCAG 62
Qy 21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40
Db 63 GCGGCTCTCTGAGTAGAAGCAGAGGTCCCGGAGCAGCAGGCTCTCTCGAAGCGGAG 122
Qy 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGlnGluLeuSer 60
Db 123 GGGGCTCAAGCCAGAACCGCTCAGTCTGGGGCCCTTCGTGTGTCTCTGAGGAGCTGAGC 182
Qy 61 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGly 80
Db 183 CGCCAACTGGAGACATACTGAGCACATACTGTGTGGACATAACCCAGGGGGGCCCGCC 242
Qy 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGlnLysSerArgThrTyr 100
Db 243 GAGGATGGGCGACAGCGCTGAGCGGCTGAACCCGGAAGATCGCAGAGAGTCCCGGAGCT 302
Qy 101 ValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGluLysGluProSer 120
Db 303 GTGGCAAGAGATGGGAGCGCTGAACCACTCCAGTAGTCTATGGAGAGAGGACCCCTCC 362
Qy 121 LysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHis 140
Db 363 AAGCGGGATCCAAACACACAGAGATCCGCGAGATGACGAGTCCGAGACCGAGACCAT 422
Qy 141 ArgArgProGlnGlnLysLysAlaLysGlyLeuLysGluIleThrLeuMet 160
Db 423 CGAAGCCACAGAGAGAGAAAGCCAAAGGTTGGGGAAGAGATCAGTGTCTGATG 482
Qy 161 GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaAlaLeuCysLysTyr 180
Db 483 CAGACATTGAATACTCTGAGTACCCACAGAGAGAGAGTGGCTGCTGTGTGCAAGAGTAT 542
Qy 181 AlaGluLeuLeuGluLysIleArgAsnSerGlnLysGlnMetLysLeuGlnLysLys 200
Db 543 GCTGAACCTGCTGAGGAGCAGCCGGAATTCACAGAGCAGATGAAGCTCTTACAGAAAAA 602
Qy 201 GlnSerGlnLeuValGlnGlnLysAspHisLeuArgGlyGluHisSerLysAlaValLeu 220
Db 603 CAGAGCCAGCTGGTGCAAGAGAGAGACCACTGCGCGGTGAGCAGCAGAGCGGCTCTG 662
Qy 221 AlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLys 240
Db 663 GCGCGCAGCAAGCTTGAGAGCCTATGCGGTGAGTGCAGCGGCACACACCGCTCCCTCAAG 722
Qy 241 GluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHis 260
Db 723 GAAGAAGGTGTGAGCGGGCCCGGAGGAGAGAGAGCGCAAGAGGTGAGCTCTCGCAC 782
Qy 261 PheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSer 280
Db 783 TTCCAGGTGACACTGAATGACATTGAGCTGAGATGGAACACAGCAATGAGCGCAACTCC 842
Qy 281 LysLeuArgGlnGlnAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGlnTyr 300
Db 843 AAGCTCGCCCAAGAGAACATGAGCTGGCTGAGAGGCTCAAGAAGCTGATTGAGCAGTAT 902
Qy 301 GluLeuArgGluGlnHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeu 320
Db 903 GAGCTCGGAGGAGGATATTCACAAAGTCTTCAACACAAAGGACCTTACACAGCAGCTG 962
```

QY 321 ValAspAlaLysLeuGlnAlaGlnGluMetLeuLysGluAlaGluAArgHisGln 340
 DB GTGGATGCAAGCTCCAGACGCCAGAGATGCTNAAGAGGAGGAGAGAGCGGCACAG 1022
 QY 341 ArgGluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMet 360
 DB CGGAGAAAGGATTTCTCTCTAAAGAGGAGCAGTAGATCCAGAGGATGTGTGAGCTGATG 1082
 QY 361 LysGlnGlnGluThrHisLysGlnGlnLeuAlaLeuThrGluLysPheGluGlu 380
 DB AAGCAGCAAGAGACCCACTGAGCAGCAGCTTCCCTATACAGAGAGATTGTGAGAG 1142
 QY 381 PheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGlu 400
 DB TTCCAGAACACACTTTTCCAAAGCAGCAGGAGTATTCACCACTTCAAGCAGGAGATGAA 1202
 QY 401 LysMetThrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 420
 DB AAGATGCTAAGAGATCAAGAGCTGGAGAAAGAAACCACTATGACCGGTCCCGGTGG 1262
 QY 421 GluSerSerAsnLysAlaLeuLeuMetAlaGluGluLysThrValArgAspLysGlu 440
 DB GAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1322
 QY 441 LeuGluGlnLysValLysLysLysLysLysLysLysLysLysLysLysLysLysLys 460
 DB CTGAGGGGCTGTCAGGTAAATCAACGGCTGGAGAGCTGTGCGGGCAGCTCAGACA 1382
 QY 461 GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnLysSerLeu 480
 DB GAGCCCATGACCTGAAACAGAGGTACAGACCTGAGTGTGTGTGCGCCAGGGCTCCCTC 1442
 QY 481 ThrAspSerGlyProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerPro 500
 DB ACTGACAGTGGCCCTGAGAGAGGCGCAGAGGGCTGGGGCTCAAGCAGCAGCTCCCCC 1502
 QY 501 ArgValThrGluAlaProCysTyProGlyAlaProSerThrGluAlaSerGlyGlnThr 520
 DB AGGGTCACAGAGCGCTTGTCTACCCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1562
 QY 521 GlyProGlnGluProThrSerAlaArgAla 530
 DB GGGCTCTCAGAGCCCACTCCGCCAGGCCC 1592

RESULT 7

US-09-976-740-46
 ; Sequence 46, Application US/09976740
 ; Publication No. US20020194633A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lees, Ann M.
 ; APPLICANT: Lees, Robert S.
 ; APPLICANT: Law, Simon W.
 ; APPLICANT: Arjona, Anibal A.
 ; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
 ; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
 ; TITLE OF INVENTION: ATHEROSCLEROSIS
 ; FILE REFERENCE: 10797-004001
 ; CURRENT APPLICATION NUMBER: US/09/976,740
 ; CURRENT FILING DATE: 2001-10-12
 ; PRIOR APPLICATION NUMBER: 09/616,289
 ; PRIOR FILING DATE: 2000-07-14
 ; PRIOR APPLICATION NUMBER: US 08/979,608
 ; PRIOR FILING DATE: 1997-11-26
 ; PRIOR APPLICATION NUMBER: US 60/031,930
 ; PRIOR FILING DATE: 1996-11-27
 ; PRIOR APPLICATION NUMBER: US 60/048,547
 ; PRIOR FILING DATE: 1997-06-03
 ; NUMBER OF SEQ ID NOS: 53
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 46
 ; LENGTH: 1638
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens

; FEATURE:
 ; NAME/KEY: CDS.
 ; LOCATION: (1)...(1638)
 US-09-976-740-46
 Alignment Scores:
 Pred. No.: 1,066-209 Length: 1638
 Score: 2693.00 Matches: 529
 Percent Similarity: 99.81% Conservative: 0
 Best Local Similarity: 99.81% Mismatches: 1
 Query Match: 99.67% Indels: 0
 DB: 9 Gaps: 0

US-10-023-523-8 (1-530) x US-09-976-740-46 (1-1638)

QY 1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 20
 DB 49 AAAGCAGCCAGGACACACCGAGAGAGCCAGGAGGAGCCAGAGCGGCCAGCCAG 108
 QY 21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40
 DB 109 GGGCTCTCTGCTAGTAGAGCAGAGGTCCTCCGCGCAGCAGCGCTCTCTCGAAGCGCGAG 168
 QY 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 60
 DB 169 GGTGCTCAAGCCAGAGACCGCTCAGTCTGGGGCCCTTGTGATGTTCTCTGAGAGCTGAGC 228
 QY 61 ArgGlnLeuGluAspLysLeuSerThrTyCysValAspAsnAsnGlnGlyGlyProGly 80
 DB 229 CGCCAACTCGAAGACATCTAGAGCAGCATCTGTGTGGACATAAACCAGGGGGGCCCGCCG 288
 QY 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTy 100
 DB 289 GAGGATGGGGCAGCAGGCTGAGCGCTGAACCCAGAGATGCGAGAGAGTCCCGGACCTAT 348
 QY 101 ValAlaArgAsnGlyGluProGluProThrProValValTyGlyGluLysGluProSer 120
 DB 349 GTGGCAGGAATGGGGAGCTGAACCACTCCAGTAGTCAATGGAGAGAGAACCTCC 408
 QY 121 LysGlyAspProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAspHis 140
 DB 409 AAGGGGGATCCAAACACAGAGAGATCCGGCAGAGTGCAGAGTCCGAGACCGAGCAT 468
 QY 141 ArgArgProGlnGluLysLysLysAlaLysGlyLeuGlyLysGluLeuThrLeuLeuMet 160
 DB 469 CGAAGCCACAG 528
 QY 161 GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysTy 180
 DB 529 CAGACATTGAATACTCTGAGTACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 588
 QY 181 AlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuGluLysLys 200
 DB 589 GCTGAATCTCTGAG 648
 QY 201 GlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeu 220
 DB 649 CAGAGCCAGCTGGTGCAG 708
 QY 221 AlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLys 240
 DB 709 GCCCGCAGCAAGCTTGAAGAGCTATGCCGTGAGTGCAGCGGAGAGAGAGAGAGAGAG 768
 QY 241 GluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHis 260
 DB 769 GAAGAAGTGTGCAGCGGGCCCGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 828
 QY 261 PheGlnValThrLeuAsnAspLysGlnLeuGlnMetGluGlnHisAsnGluArgAsnSer 280
 DB 829 TTCCAGGTGACACTGAATGACATTCAGTGTGAGATGGAAACAGCAGCAATGAGCGCACTCC 888
 QY 281 LysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuGluGlnTy 300


```

Db 889 AAGCTGCCCAAGAGCAATCGAGCTGGCTGAGAGGCTCAAGAGCTGATTGACAGTAT 948
QY 301 GluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeu 320
Db 949 GAGTGGCCGAGGAGCATATCGACAAAGTCTTCAACACAGGACCTACACAGCAGCTG 1008
QY 321 ValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGln 340
Db 1009 GTGATGCCAAGCTCCAGCAGCCCGAGAGTGTAAAGGAGGAGAGGCGGCACAG 1068
QY 341 ArgGluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMet 360
Db 1069 CGGAGAGAGGATTTCTCTCTCAAGAGGAGCTAGAGTCCCGAGAGGATGTGTGAGCTGATG 1128
QY 361 LysGlnGlnGlnThrHisLysGlnGlnLeuAlaLeuThrGluLysPheGluGlu 380
Db 1129 AAGCAGCAGAGACCCACCTGAACACAGCTTCCCTTATACAGAGAGTGTAGGAG 1188
QY 381 PheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGlu 400
Db 1189 TTCCAGAACACACTTCCAAAGCAGCGAGGTATTCCACCACATTCAGAGGAGATGGAA 1248
QY 401 LysMetThrLysLysIleLysLysLeuGluLysGlnThrThrMetTyrArgSerArgTyr 420
Db 1249 AAGATGACTAAGAGAGATCAAGAGCTGGAGAAAGAAACCCATGTATCCCGTCCCGGTGG 1308
QY 421 GluSerSerAsnLysAlaLeuLeuGluMetAlaGluLysThrValArgAspLysGlu 440
Db 1309 GAGAGCAGCAACAGGCTCTTGTAGTGGCTGAGGAGAAACAGTCCGGGATTAAGAA 1368
QY 441 LeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThr 460
Db 1369 CTGAGAGGCTGCGAGGTAAATCCACCGCTGGAGAGCTGTGCCGGGCACTGCAGACA 1428
QY 461 GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlnGlySerLeu 480
Db 1429 GAGCGCAATGACCTGAACAGAGGATACAGACCTGAGTGTGTGGCCAGGCTCCCTC 1488
QY 481 ThrAspSerGlyProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerPro 500
Db 1489 ACTCACAGTGGCTCGAGAGGAGCCAGAGGGGCTGGGGCTCAAGCACCCAGCTCCCC 1548
QY 501 ArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThr 520
Db 1549 AGGTCACAGAGGCGCTTCTACCCAGGAGCAGGAGCAGAGCAGAGCATCAGGCCAGACT 1608
QY 521 GlyProGlnGluProThrSerAlaArgAla 530
Db 1609 GGGCTCAAGAGCCACCTCCGCCAGGAGGCC 1638

```

RESULT 8

```

US-10-671-242-46
; Sequence 46, Application US/10671242
; Publication No. US2004004049A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/671,242
; CURRENT FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27

```

```

; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 1638
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1638)
; US-10-671-242-46

```

Alignment Scores:

```

Pred. No.: 1,06e-209 Length: 1638
Score: 2693.00 Matches: 529
Percent Similarity: 99.81% Conservative: 0
Best Local Similarity: 99.81% Mismatches: 1
Query Match: 99.67% Indels: 0
DB: 13 Gaps: 0

```

US-10-023-523-8 (1-530) X US-10-671-242-46 (1-1638)

```

QY 1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 20
Db 49 AAAAGCAGCCAGGACACACCGGAGCAGCAGCCGAGGAGCCAGAGCGGCCAGCCAG 108
QY 21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40
Db 109 GCGGCTCTTCAGTAGTAAAGCAGAGGTCCCGGACAGCCAGCGCTCTCGGAAGCCGGAG 168
QY 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 60
Db 169 GGTGCTCAAGCCAGAACCGCTCAGTCTGGGGCCCTTCGTGTGTCTCTCAGGAGCTGAGC 228
QY 61 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGly 80
Db 229 CGCCAACTGGAAGACATATCTGAGCACATCTGTGTGCAATAACACAGGGGGGCCCGGC 288
QY 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
Db 289 GAGGATGGGCAACAGGGTGAGCGGCTGAACCCGAGATGACAGAGAGTCCCGACCTAT 348
QY 101 ValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGluGluProSer 120
Db 349 GTGGCAAGGATGGGAGCGCTGAACCAACTCCAGTAGTCAATGGAGAGAGAACCCCTCC 408
QY 121 LysGlyAspProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgHis 140
Db 409 AAGGGGGATCCAAACACAGAGAGATCCGAGAGTCCGAGAGTCCGAGAGCCGAGACCAT 468
QY 141 ArgArgProGlnGluLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMet 160
Db 469 CGAAGGCCACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 528
QY 161 GlnThrLeuAsnThrLeuSerThrProGluLysLeuAlaAlaLeuCysLysLysTyr 180
Db 529 CAGACATTGAATACCTCTCAGTACCCCGAGAGAGAGCTGGCTGTCTCTGCAAGAAGTAT 588
QY 181 AlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLys 200
Db 589 GGTGACTGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 648
QY 201 GlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeu 220
Db 649 CAGAGCCAGCTGTGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 708
QY 221 AlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLys 240
Db 709 GCGCCAGCAGAGCTTGAAGAGCTATGCGCTGAGCTGAGCGGACCAACCGCTCCCTCAG 768
QY 241 GluGluGlyValGlnArgAlaArgGluGluLysArgLysGluValThrSerHis 260

```


Db 649 CAGAGCCAGCTGGTGCAGAGAGAGACACCTGCGGGTGAGCACAGCAAGCCGCTCCTG 708
QY 221 AlaArgSerLysLeuGlnSerLysCysArgGluLeuGlnArgHisAsnArgSerLeuLys 240
Db 709 GCGCCAGCAGCAAGCTTTGAGAGCCTATCGGTGAGCTGCAGCGGCACAAACCGCTCCCTCAAG 768
QY 241 GluGluGlyValGlnArgAlaArgGluGluGluGluLysArgLysGluValThrSerHis 260
Db 769 GAAGAGGTGTCCAGCGGGCCCGGAGGAGGAGAGAGCGCAAGAGGTGACCTCGGCAC 828
QY 261 PheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSer 280
Db 829 TTCCAGGTGCACACTGAATGCATTCAGCTGCAGATGGAACAGCACAAATGAGCGCACTCC 888
QY 281 LysLeuArgGlnGlnLysMetLysGluGlnArgLysLysLeuLysGluGlnTyr 300
Db 889 AAGCTGCGCCAGAGAAATGAGCTGCTGAGAGCTCAAGAGCTGATTGAGCAGTAT 948
QY 301 GluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeu 320
Db 949 GAGCTGCGCGAGGACATATCGCAAAAGTCTTCAACACACAGGACCTACACAGCAGCTG 1008
QY 321 ValAspAlaLysLeuGlnGlnAlaGlnGluMetLysLysGluAlaGluArgHisGln 340
Db 1009 GTGGATGCAAGCTCCAGCAGCCAGGAGATGCTTAAAGGAGGAGAGAGCGGCACCCAG 1068
QY 341 ArgGluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMet 360
Db 1069 CGGAGAGAGGATTTCTCTGAAAGAGGAGGAGTAGAGTCCAGAGGATGTGAGCTGATG 1128
QY 361 LysGlnGlnGluThrHisLysLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGlu 380
Db 1129 AAGCAGCAAGAGACCCACTGAAGCAACAGCTTGCCCTATACACAGAGAAGTTTGAGGAG 1188
QY 381 PheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGlu 400
Db 1189 TTCAGAAACACACTTTCAGAGAGGAGGAGGATTTCCACCATTCAGCAGGAGATGGAA 1248
QY 401 LysMetThrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 420
Db 1249 AAGATGACTAAGAGATCAAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1308
QY 421 GluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGlu 440
Db 1309 GAGAGCAGCAACAAGGCCCTGCTTGAGATGGCTGAGGAGAGAGAGAGAGAGAGAGAGAG 1368
QY 441 LeuGluGlyLeuGlnValLysLysLysLysLysLysLysLysLysLysLysLysLys 460
Db 1369 CTGGAGGCTGCGAGTAAATCCACGGCTGAGAGAGCTGTGCGGGGCACTGCAGACA 1428
QY 461 GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlnGlnSerLeu 480
Db 1429 GAGCGCAATGACCTGAACAAGAGGATACAGGACCTGAGTGTGCTGGCCAGGGCTCCCTC 1488
QY 481 ThrAspSerGlyProGluArgArgProGluGlyProGluGlyProGluGlyProGlu 500
Db 1489 ACTGACGTGCGCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1548
QY 501 ArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThr 520
Db 1549 AGGGTCACAGAAGCGCTTGTATCCAGGAGCACCAGAGCAGAGAGCATCAGGCGCAGCT 1608
QY 521 GlyProGlnGluProThrSerAlaArgAla 530
Db 1609 GGGGCTCAAGAGGCCACCTCCGCGAGGGGCC 1638

RESULT 10

US-10-023-523-46

; Sequence 46, Application US/10023523

; Publication No. US20020152485A1

; GENERAL INFORMATION:

; APPLICANT: Lees, Ann M.

; APPLICANT: Lees, Robert S.

; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,523
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 1638
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1638)
US-10-023-523-46

Alignment Scores:

Pred. No.: 1,06e-209 Length: 1638
Score: 2693.00 Matches: 529
Percent Similarity: 99.81% Conservative: 0
Best Local Similarity: 99.81% Mismatches: 1
Query Match: 19.67% Indels: 0
DB: 14 Gaps: 0

US-10-023-523-8 (1-530) x US-10-023-523-46 (1-1638)

QY 1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyValaGlnGluArgProSerGln 20
Db 49 AAAAGCAGCCAGGACAAACCGGAAGCAGGACCCGAGGAGCCAGGAGCGGCCAGCCAG 108
QY 21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40
Db 109 GCGGCTCTCTGCTAGCAAGCAGAGGTCCTCCGCGCAGCAGCCAGGCTCTCGAAGCCGAG 168
QY 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 60
Db 169 GGTGCTCAAGCCAGAAACGGCTCAGTCTGGGGCCCTTCGTGATGTCTCTGAGGAGCTGAGC 228
QY 61 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnGlnGlyGlyProGly 80
Db 229 CGCCAACTGGAGACATACCTAGCACATCTGTGTGGACAAATACCCAGGGGGGCCCGGC 288
QY 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
Db 289 GAGGATGGGGCAGCAGCGGTGAGCGGTGAACCCGAGATCCAGAGAGTCCCGGAGCTAT 348
QY 101 ValAlaArgAsnGlyGluProGluProThrProValValTyrGlyLysGluProSer 120
Db 349 GTGGCAAGGAATGGGAGGCTGAACCACTCCAGTAGTCAATGGAGAGAGAACCTCC 408
QY 121 LysGlyAspProAsnThrGluGluLeuArgGlnSerAspGluValcLysArgAspHis 140
Db 409 AAGGGGATCCAAACACACAGAGATCCGGCAGAGTACGAGGTCCGAGACCCGAGACCAT 468
QY 141 ArgArgProGlnGluLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 160
Db 469 CGAAGGCCACAG 528
QY 161 GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuLysLysTyr 180

529 DB CAGACATTGAATACCTCTGAGTACCCAGAGAGAGCTGGCTGCTCTGTGCAAGAGTAT 588
 181 QY AlaGluLeuLeuGluGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLys 200
 589 DB GCTGAACTGCTGGAGAGACACCGAATTCACAGAGACAGATGAAGTCTCTACAGAAAAG 648
 201 QY GlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeu 220
 649 DB CAGAGCCAGCTGGTCAAGAGAGACACCTGGCGGTGAGCAGCAGCAAGGCCGCTCCTG 708
 221 QY AlaArgSerLysLeuGlnSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLys 240
 709 DB GCGCCAGCAGACTTGAGAGCTATGCGTGAAGCTGAGCGGCGCAGCAACCGCTCCCTCAAG 768
 241 QY GluGluGlyValGlnArgAlaArgGluGluGluLysArgLysArgLysGluValThrSerHis 260
 769 DB GAAGAAGGTGTGACGGCGCCGAGGAGGAGAGAGCGCAAGGAGGTGACCTCGCAC 828
 261 QY PheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSer 280
 829 DB TTCCAGGTGACACTGAATGCAATTCAGCTGCAGATGGAACAGCAGCATGAGCGCACTCC 888
 281 QY LysLeuArgGlnGlnAsnMetGluLeuAlaGluArgLysLysLysLeuIleGluGlnTyr 300
 889 DB AAGTGGCGCAAGAGACAATGAGCTGCTGAGAGGCTCAAGAAAGCTGATTGAGCAGTAT 948
 301 QY GluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeu 320
 949 DB GAGCTGCGCGAGGAGCATATGCAAAAGTCTCAACACAGAGGACCTCAACAGCAGCTG 1008
 321 QY ValAspAlaLysLeuGlnGlnAlaGlnMetLeuLysGluAlaGluGluArgHisGln 340
 1009 DB GTGGATGCCAAGCTCCAGAGCCCGAGGAGATGCTAAGAGGCGCAGAGAGCGCAGCAG 1068
 341 QY ArgGluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMet 360
 1069 DB CGGAGAGAGGATTTCTCTCTGAAGAGGAGTAGAGTCCAGAGGATGTGAGCTGATG 1128
 361 QY LysGlnGlnGluThrHisLysLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGlu 380
 1129 DB AAGCAGCAGAGACCCACCTGAGCAACAGCTTGCCCTATACAGAGAGAGTTTGAGGAG 1188
 381 QY PheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGlu 400
 1189 DB TTCCAGAACACACTTCCAAAAGCAGCAGAGGTATTCAACACATTCAAGCAGAGAGTGA 1248
 401 QY LysMetThrLysLysLysLysLysLysGluThrThrMetTyrArgSerArgTyr 420
 1249 DB AAGATGACTAAGAGATCAAGAGCTGGAGAAAGAACCCATGATCGGTCCTCCGTTGG 1308
 421 QY GluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGlu 440
 1309 DB GAGAGCAGCAACAGGCCCTGCTGAGATGGCTGAGGAGAAACAGTCCGGGATAAGAA 1368
 441 QY LeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLysCysArgAlaLeuGlnThr 460
 1369 DB CTGGAGGGCTCGCAGGTAAAAATCCAAAGCTGGAGAGAGCTGTGCGGCACTGCAGACA 1428
 461 QY GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeu 480
 1429 DB GAGCGCAATGACTGACAGAGAGGGTACAGGACCTGAGTGTGTTGGTGGCCAGGGCTCC 1488
 481 QY ThrAspSerGlyProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerPro 500
 1489 DB ACTGACAGTGGCCCTGAGAGAGGCGCAGAGGGGCTGGGGCTTCAAGCACCCAGCTCCC 1548
 501 QY ArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThr 520
 1549 DB AGGTTCAGAGAGCGCTTGTCTACCCAGGAGGACCCGAGCAGCAGAGCATCAGGCCAG 1608
 521 QY GlyProGlnGluProThrSerAlaArgAla 530
 1609 DB GGGCCTCAAGAGAGCCACCTCCCGCAGGGCC 1638

RESULT 11

US-10-616-187-46.
 ; Sequence 46, Application US/10616187
 ; Publication No. US20040013668A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lees, Ann M.
 ; APPLICANT: Lees, Robert S.
 ; APPLICANT: Law, Simon W.
 ; APPLICANT: Arjona, Anibal A.
 ; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
 ; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
 ; TITLE OF INVENTION: ATHEROSCLEROSIS
 ; FILE REFERENCE: 10797-004001
 ; CURRENT APPLICATION NUMBER: US/10/616,187
 ; CURRENT FILING DATE: 2003-07-09
 ; PRIOR APPLICATION NUMBER: US/09/616,289
 ; PRIOR FILING DATE: 2000-07-14
 ; PRIOR APPLICATION NUMBER: US/09/517,849
 ; PRIOR FILING DATE: 2000-03-02
 ; PRIOR APPLICATION NUMBER: US/08/979,608
 ; PRIOR FILING DATE: 1997-11-26
 ; PRIOR APPLICATION NUMBER: US/60/031,930
 ; PRIOR FILING DATE: 1996-11-27
 ; PRIOR APPLICATION NUMBER: US/60/048,547
 ; PRIOR FILING DATE: 1997-06-03
 ; NUMBER OF SEQ ID NOS: 53
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 46
 ; LENGTH: 1638

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (1)...(1638)

US-10-616-187-46

Alignment Scores:

Pred. No.: 1,066-209 Length: 1638

Score: 2693.00 Matches: 529

Percent Similarity: 99.81% Conservative: 0

Best Local Similarity: 99.81% Mismatches: 1

Query Match: 99.67% Indels: 0

DB: 16 Gaps: 0

US-10-023-523-8 (1-530) x US-10-616-187-46 (1-1638)

QY 1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 20
 DB 49 AAAAGCAGCCAGCAGACACCGGAGCAGGAGCCAGGAGCCAGGAGCGGCCAGCCAG 108
 QY 21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40
 DB 109 GCGGCTCTCTCAGTAGAAGCAGAGGTCGCGCAGCAGCAGGCTCTCTCGGAGCCGGAG 168
 QY 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 60
 DB 169 GTGTCTCAACGACAGACGGCTCAGTCTGGGGCCCTTCGTGATGCTCTCTGAGGAGCTGAGC 228
 QY 61 ArgGlnLeuGluAspLysLeuSerThrTyrCysValAspAsnAsnGlnGlyProGly 80
 DB 229 CGCAACTGGAAGACATCTAGCAGCATACTGTGTGACAAATAACCCAGGGGGGGCCCGGC 288
 QY 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
 DB 289 GAGGATGGGACACAGGCTGAGCGGCTGAACCCAGAGATGCAGAGAGTCCCGGACCTAT 348
 QY 101 ValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGluLysGluProSer 120
 DB 349 GTGCAAGGAATGGGAGCGCTGAACCACTCCAGTAGTCAATGAGAGAGGAAGAACCTCC 408
 QY 121 LysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHis 140

Db 409 AAGGGGATCCAAACACAGAGAGATCCGGCAGAGTGCAGAGTCCGAGACCGAGACCAT 468
Qy 141 ArgArgProGlnGluLysLysAlaLysGlyLeuGlyLysGluLeuThrLeuLeuMet 160
Db 469 CGAAGGCCACAG 528
Qy 161 GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysTyr 180
Db 529 CAGCATGTGAATCTCTAGTACCCCAAGAGAGAGAGCTGGCTCTCTGTCAAGAGATAT 588
Qy 181 AlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLys 200
Db 589 GCTGAATGCTGGAG 648
Qy 201 GlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeu 220
Db 649 CAGAGCAGCTGGTGCAG 708
Qy 221 AlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLys 240
Db 709 GCCCGCAGCAAGCTTGAGAGCTATGCGGTAGCTGCAGCGGCACACCGCTCCCTCAAG 768
Qy 241 GluGluGlyValGlnAArgGluGluGluGluLysArgLysGluValThrSerHis 260
Db 769 GAAGAAGGTGTGCAGCGCGGCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 828
Qy 261 PheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSer 280
Db 829 TTCAGGTGACACTGAATGACATTCAGCTGCAGATGGAACAGACACATGAGCGCACTCC 888
Qy 281 LysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuLeuGluGlnTyr 300
Db 889 AAGCTCGCCCAAGAGACATGGAGCTGGCTGAGAGGCTCAAGAAGCTGATTGAGCAGTAT 948
Qy 301 GluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeu 320
Db 949 GAGCTGCGCAGAGAGAGATATCGAACAAGTCTTCAACACAGAGAGAGAGAGAGAGAG 1008
Qy 321 ValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGln 340
Db 1009 GTGATGCCAAGCTCCAGCGCCCGAGAGATGCTTAAGGAGGAGAGAGAGAGAGAGAG 1068
Qy 341 ArgGluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMet 360
Db 1069 CGGAGAGAGATTTCTCTCTGAAGAGCGCAGTAGAGTCCAGAGGATGTGTGAGCTGATG 1128
Qy 361 LysGlnGlnThrHisLysLysGlnLeuAlaLeuTyrThrGluLysPheGluGlu 380
Db 1129 AAGCAGCAAGAGACCCACCTGAGCAACAGCTTGCCCTATACAGAGAGATTTGAGGAG 1188
Qy 381 PheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGlu 400
Db 1189 TTCAGAAACACACTTCCAAAGCAGCGAGGTATTCACCACTTCAAGCAGGAGATGGAA 1248
Qy 401 LysMetThrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 420
Db 1249 AAGATGACTAAGAAGATCAAGAAGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1308
Qy 421 GluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGlu 440
Db 1309 GAGAGCAGCAACAGAGCCCTGCTTGAGATGGCTGAGGAGAGAGAGAGAGAGAGAGAG 1368
Qy 441 LeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLysCysArgAlaLeuGlnThr 460
Db 1369 CTGAGAGGCTCTGAGAGTAAATAATCCACCGCTGGAGAGAGAGAGAGAGAGAGAGAG 1428
Qy 461 GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeu 480
Db 1429 GAGCGCATGACCTGAACAG 1488
Qy 481 ThrAspSerGlyProGluAArgProGluGlyProGlyAlaGlnAlaProSerSerPro 500
Db 1489 ACTCACAGTGGCCCTGAG 1548

Qy 501 ArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThr 520
Db 1549 AGGGTCCACAG 1608
Qy 521 GlyProGlnGluProThrSerAlaArgAla 530
Db 1609 GGGCTCAAG 1638

RESULT 12

US-10-276-774-784
; Sequence 784, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Fang, Y, Tom et al
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 784
; LENGTH: 2523
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-276-774-784

Alignment Scores:
Pred. No.: 1,77e-209 Length: 2523
Score: 2693.00 Matches: 529
Percent Similarity: 99.81% Conservative: 0
Best Local Similarity: 99.81% Mismatches: 1
Query Match: 99.67% Indels: 0
DB: 13 Gaps: 0

US-10-023-523-8 (1-530) x US-10-276-774-784 (1-2523)

Qy 1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 20
Db 167 AAAAGCAGCCAGGACCAACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 226
Qy 21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40
Db 227 GCGGCTCCTGCACTAG 286
Qy 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 60
Db 287 GGGGCTCAAGCCAG 346
Qy 61 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGly 80
Db 347 CGCCAACTGGAAGACATCTGAGCAGATCTGTGGACATTAACCAACAGAGAGAGAGAG 406
Qy 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
Db 407 GAGGATGGGCAACAGGTTGAGCGGCTGAACCCGAGAGATGACAGAGAGTCCCGGACCTAT 466
Qy 101 ValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGluLysGluProSer 120
Db 467 GTGGCAAGAAATGGGAG 526
Qy 121 LysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHis 140
Db 527 AAGGGGATCCAAACACAGAGAGATCCGAGAGATGAGAGAGTCCGAGAGAGAGAGAG 586
Qy 141 ArgArgProGlnGluLysLysLysLysLysLysLysLysLysLysLysLysLysLys 160
Db 587 CGAAGGCCACAG 646

QY 161 GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysLys 180
DB 647 CAGACATTGATCTCTGAGTACCCAGAGAGAGCTGGCTGCTCTGTGCAAGAGTAT 706
QY 181 AlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLys 200
DB 707 GCTGAACCTGCTGGAGGAGCACCGGAATTCACAGAGAGAGATGAAGCTCTACAGAAAAG 766
QY 201 GlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeu 220
DB 767 CAGACCCAGCTGTGTCAGAGAGAGACCACTGGCGGTGAGCAGCAGAGCGCTCTCTG 826
QY 221 AlaArgSerLysLeuGluSerLysCysArgGluLeuGlnArgHisAsnArgSerLeuLys 240
DB 827 GCCCGCAGCAGCTTGAGAGCTATGCTGCTGAGCTGAGCGCGCACAACCGCTCCTCAAG 886
QY 241 GlnGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHis 260
DB 887 GAAGAAGGTGTGACGGGCGCGGAGGAGAGAGCGCAAGAGAGGTGACCTCCAC 946
QY 261 PheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSer 280
DB 947 TTCAGGTGACATGATGACATTCAGCTGCAGATGGAACAGCACAATGAGCGCACTCC 1006
QY 281 LysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuLleGluGlnThr 300
DB 1007 AAGCTGCCCAAGAGAAATGAGCTGGCTGAGAGGCTCAAGAAGCTGATTGACAGTAT 1066
QY 301 GluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeu 320
DB 1067 GAGCTGCCGAGGAGCATATCGCAAACTCTTCAACACAGAGACTACACACAGCTG 1126
QY 321 ValAspAlaLysLeuGlnGlnAlaGlnGluMetLysLysGluAlaGluGluArgHisGln 340
DB 1127 GTGGATGCCAAGCTCCAGCAGCGCCAGAGAGATGTAAGAGGAGGAGAGAGCGCCACAG 1186
QY 341 ArgGluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMet 360
DB 1187 CCGGAGAAGGATTTCTCTGAAGAGAGCAGTAGAGTCCACAGAGAGTGTGAGCTGATG 1246
QY 361 LysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuThrGluLysPheGluGlu 380
DB 1247 AAGCAGCAAGAGAGACCCACTGAGCAACAGCTTGCCCTATACAGAGAAATTTGAGGAG 1306
QY 381 PheGlnAsnThrLeuSerLysSerSerGluValPheThrPheLysGlnGluMetGlu 400
DB 1307 TTCAGAACACACTTTCAAAGAGCAGGAGGTATTCACCAATTCAGAGGAGATGAA 1366
QY 401 LysMetThrLysLysLysLysLeuGluLysGluThrThrMetThrArgSerArgTyr 420
DB 1367 AAGATGACTAGAAGATCAAGAGCTGAGAAAGAACCAACCATGTACCGTCCCGGTGG 1426
QY 421 GluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGlu 440
DB 1427 CAGAGCAGCAACAGCGCCCTGTTGAGATGCTGAGGAGAAACAGTCCCGGATTAAGAA 1486
QY 441 LeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThr 460
DB 1487 CTGGAGGAGGCTTCAGTAAATCAACGGCTGAGAGAGCTGTCCCGGCTCTGACAGCA 1546
QY 461 GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeu 480
DB 1547 GAGCGCAATGACCTCAACAGAGGATCAGGACCTGAGTGTGTGGCCAGGGGTCTCCCTC 1606
QY 481 ThrAspSerGlyProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerPro 500
DB 1607 ACTGACATGGCCCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1666
QY 501 ArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThr 520
DB 1667 AGGGTCACAGAGCGCTTGTACCCAGAGAGCAGCAGCAGCAGAGATCAGGCGCAGACT 1726

QY 521 GlyProGlnGluProThrSerAlaArgAla 530
DB 1727 GGGCCTCAAGAGCCCACTCCGCCAGGCC 1756

RESULT 13

US-09-962-055-14

; Sequence 14, Application US/09962055

; Patent No. US20020052033A1

; GENERAL INFORMATION:

; APPLICANT: Lees, Ann M.

; Law, Robert S.

; Arizona, Anibal A.

; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN

; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND

; TREATING ATHEROSCLEROSIS

; NUMBER OF SEQUENCES: 42

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02110-2804

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/962,055

; FILING DATE: 24-Sep-2001

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/979,608

; FILING DATE: 26-NOV-1997

; APPLICATION NUMBER: US 60/031,930

; FILING DATE: 27-NOV-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Myers, Louis

; REGISTRATION NUMBER: 35,965

; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617/542-5070

; TELEFAX: 617/542-8906

; INFORMATION FOR SEQ ID NO: 14:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4722 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; FEATURE:

; NAME/KEY: Coding Sequence

; LOCATION: 61...1731

; SEQUENCE DESCRIPTION: SEQ ID NO: 14:

US-09-962-055-14

Alignment Scores:

Pred. No.: 1,18e-190 Length: 4722

Score: 2465.50 Matches: 492

Percent Similarity: 92.99% Conservative: 12

Best Local Similarity: 90.77% Mismatches: 25

Query Match: 91.25% Indels: 13

DB: 9 Gaps: 3

US-10-023-523-8 (1-530) x US-09-962-055-14 (1-4722)

QY 1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 20

DB 109 AAAAGCAGCCCGGACAGCCGAGCAGAGCGAGGAGGCCCGGCGGCGGCGGCGG 168

QY 21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40

DB 169 CCGGGCCCCCGCCGAGAGCCGAAGGT---GCCAGCAGCCAGGCTCCCGGGGAGCGCGGAG 225

QY 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 60
 Db 226 CGGGCTCAGCCAAACTGCTCAGCTGGGGCGCTGTGTGATGTCTCTGAGGAGCTGAGC 285
 QY 61 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAenGlnGlyGlyProGly 80
 Db 286 CGCCAGTGGAGACATACTCAGTACATACTGTGTGACACCAACACAGGGGGCCCCGGGT 345
 QY 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
 Db 346 GAGGATGGGTTCAGGGTGGAGCCCTGAACTGAAGATCGAGAGAGTCTCGCGCTAT 405
 QY 101 ValAlaArgAsnGlyGluProGluPro---ThrProValValTyrGlyGluLysGluPro 119
 Db 406 GTGCAAGGAATGGGAGCGGAGCGGACCGGACCCAGTATGATGCGCGAGAGAGACC 465
 QY 120 SerLysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAsp 139
 Db 466 TCCAAGCAGAGCGGGCGGACCGGAGAGATCCGACGAGCGATGAGGTGCGAGACCGGAGC 525
 QY 140 HisArgArgProGlnGluLysLysLysAlaLysGlyLysGluIleThrLeuLeu 159
 Db 526 CACCGAGGCCACAGGAAAGAGAGAGCCAGGCTCTGGAAAGAGATCAGCGTGTG 585
 QY 160 MetGlnThrLeuAsnThrLeuSerThrProGluLysLeuAlaLeuLysLysLys 179
 Db 586 ATGCAGACACTGAACCGCTGAGCACCCAGAGAGAGAGCTGCGGCTCTGTGCAAGAG 645
 QY 180 TyrAlaGluLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuGlnLys 199
 Db 646 TATGCGGAAGTGTGCGAGGAGACCGGAACTCCAGAGCAGATGAAGTGTGTCGAGAG 705
 QY 200 LysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaVal 219
 Db 706 AAGCAGAGCCAGCTGTGTGAGAGAGAGACCACTCGGTGGCGAGCAGCAGAGGCCATC 765
 QY 220 LeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeu 239
 Db 766 CTGGCCCGCAGCAAGCTCGAGAGCCCTGTGCGGAGCTGCGGGGAGCAGCAACCGCTGCTC 825
 QY 240 LysGluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSer 259
 Db 826 AAGAGAGAGGTGTGCGAGGAGCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 885
 QY 260 HisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsn 279
 Db 886 CACTTCCAGATGACGTCAACGACATTCAGTGCAGATGAGCAGCAGCAGCAACGAGCGCAC 945
 QY 280 SerLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLeuIleGluGln 299
 Db 946 TCCAAGCTGCGCCAGGAGAACATGGAGCTGGCGGAGCGGCTCAAGAGAGTATTGAGCAG 1005
 QY 300 TyrGluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGln 319
 Db 1006 TAGGAGCTCGAGAGAGAGCAGACATCGACAAGTCTTCAACACAGAGATCTGAGCAGAG 1065
 QY 320 LeuValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHis 339
 Db 1066 CTGGTGGAGCGCAAGCTCCAGCAGGCGCCAGGAGATGCTGAGGAGCGAGAGCGGCAC 1125
 QY 340 GlnArgGluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeu 359
 Db 1126 CAGCGGGAGAGGACTTCTCTGAGGAGGCGGTGGAGTCCAGAGGATGTCGAGGTG 1185
 QY 360 MetLysGlnGlnThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGlu 379
 Db 1186 ATGAAGCAACAGAGAGACCCACCTGAAGCAGCAGCTTGCCTATACAGAGAGATTTGAG 1245
 QY 380 GluPheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMet 399
 Db 1246 GAGTTCAGAACACTCTTTCCAAAAGCAGCGAGGTGTTCACACATTCAACAGGAAATG 1305

QY 400 GluLysMetThrLysLysLysLysLysLeuGluLysGluThrThrMetTyrArgSerArg 419
 Db 1306 GAAAAGATCACAAGAAAGATCAAGAGCTGGAGAAAGAGACCAACCATGTACCGTTCCCGG 1365
 QY 420 TrpGluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLys 439
 Db 1366 TGGGAGAGCAGCAACAGGCCCTGCTTGAGATGGCTGAGGAGAAACACTCCGGGACAAA 1425
 QY 440 GluLeuGluGlyLeuGlnValLysLysLysLysLysLeuGluLysLeuCysArgAlaLeuGln 459
 Db 1426 GAGCTGGAGAGCGCTGCGAGGTGAATAATCCAGCGCTGGAGAGCTGTGCCGGGCACACTGCAG 1485
 QY 460 ThrGluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySer 479
 Db 1486 ACAGAGCGCAATGACTGAACAGAGGGTGCAGGACCTGAGTGGCGGTGGCCAGGGCCCC 1545
 QY 480 LeuThrAspSerGlyProGluArgArgPro----- 489
 Db 1546 GTCTCCGACAGCGGTCTCTGAGCGGAGGCCAGAGCCGCCACCACTCCAGGAGCAGGGT 1605
 QY 490 ---GluGlyProGlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyr 508
 Db 1606 GTCGAGGGGCCCCGGGGCTCAAGTACCACTCTCCAGGGGCCACAGACGCTTCTCTGCTGC 1665
 QY 509 ProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAla 528
 Db 1666 GCAGGTGCACCCAGCAGCAGAGGCGATCAGGCCAGAGGGCCCCAGGAGCCCACTGCC 1725
 QY 529 ArgAla 530
 Db 1726 ACTGCC 1731

RESULT 14
 US-09-976-740-14
 ; Sequence 14, Application US/09976740
 ; Publication No. US20020194633A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lees, Robert S.
 ; APPLICANT: Law, Simon W.
 ; APPLICANT: Arizona, Anibal A.
 ; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
 ; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
 ; TITLE OF INVENTION: ATHEROSCLEROSIS
 ; FILE REFERENCE: 10797-004001
 ; CURRENT APPLICATION NUMBER: US/09/976,740
 ; CURRENT FILING DATE: 2001-10-12
 ; PRIOR APPLICATION NUMBER: 09/616,289
 ; PRIOR FILING DATE: 2000-07-14
 ; PRIOR APPLICATION NUMBER: US 08/979,608
 ; PRIOR FILING DATE: 1997-11-26
 ; PRIOR APPLICATION NUMBER: US 60/031,930
 ; PRIOR FILING DATE: 1996-11-27
 ; PRIOR APPLICATION NUMBER: US 60/048,547
 ; PRIOR FILING DATE: 1997-06-03
 ; NUMBER OF SEQ ID NOS: 53
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 14
 ; LENGTH: 4722
 ; TYPE: DNA
 ; ORGANISM: Oryctolagus cuniculus
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (61)...(1731)
 US-09-976-740-14

Alignment Scores:
 Pred. No.: 1,18e-190 Length: 4722
 Score: 2465.50 Matches: 492
 Percent Similarity: 92.99% Conservative: 12
 Best Local Similarity: 90.77% Mismatches: 25
 Query Match: 91.25% Indels: 13
 DB: 9 Gaps: 3

US-10-023-523-8 (1-530) x US-09-976-740-14 (1-4722)

QY 1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 20
DB 109 AAAGCAGCCGGGACAGCCGGAACAGGAGCGGAGGAGCCAGGCGCGCCGCGCG 168
QY 21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40
DB 169 CCGGCCCCCGCCGAGAGCCGAGGT--GCCAGCAGCCAGGCTCCCGGAGGCGCGAG 225
QY 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 60
DB 226 GGGGCTCAAGCCAAACTGCTCAGCGCTGGGCGCTCTGTGATGCTCTGAGGAGCTGAGC 285
QY 61 ArgGlnLeuGluAspLeuSerThrTyrcysValAspAsnAsnGlnGlyGlyProGly 80
DB 286 CGCCAGTTGGAAGACATACCTCAGTACATACCTGTGTGACACACACAGGCGGCGGCT 345
QY 81 GluAspGlyAlaGlnGlyProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
DB 346 GAGGATGGGGTCCAGCGTGGCGGCTGAGCCCTGAACTGAGATGCGAGAGAGTCTCGCGCTAT 405
QY 101 ValAlaArgAsnGlyGluProGluPro--ThrProValValTyrGlyGluLysGluPro 119
DB 406 GTGGCAAGGAATGGGAGCCGAGCCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 465
QY 120 SerLysGlyAspProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAsp 139
DB 466 TCCAAAGGACAGCCGGGACGAGAGATCCGGACGAGATGAGGTGAGGTGCGGAGCCGAGAC 525
QY 140 HisArgArgProGlnGluLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeu 159
DB 526 CACCGGAGCCACAGGAAAGAGAGCCGAGGCTCTGGGAAAGAGATCAAGCTGCTG 585
QY 160 MetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuLysLysLys 179
DB 586 ATGCAGACACTGAACACGCTGAGCACCAGGAGGAGAGGAGTGGCGGCTCTGTGCAAGAG 645
QY 180 TyrAlaGluLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLys 199
DB 646 TATGGGAACTGCTCGAGGACACCGGAACTCCGAGAGCAGATGAAGTCTGTCGAGAG 705
QY 200 LysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaVal 219
DB 706 AAGCAGAGCCAGCTGTCAGAGAGAGGAGGAGCACCCTGCTGGCGGAGCAGACAGAGGCGCATC 765
QY 220 LeuAlaArgSerLysLeuGluSerLeuGluCysArgGluLeuGlnArgHisAsnArgSerLeu 239
DB 766 CTGGCCCGCAGCAAGCTCGAGAGCTGTGCGGAGCTGCGAGCGGCAACACCGCTGCTC 825
QY 240 LysGluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSer 259
DB 826 AAGGAGAGAGGTGTGAGCAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 885
QY 260 HisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsn 279
DB 886 CACTTCCAGATGACCTCAACGACATTCAGTGTGAGATGGAGCGGAGCAGCAGGAGCGCAAC 945
QY 280 SerLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLysLysLysLys 299
DB 946 TCCAAAGTCCGCGCAGGAGAAATGAGCTGCGGCGGCGGCTCAAGAGAGTGTATTGAGCAG 1005
QY 300 TyrGluLeuArgGluGluHisLysAspLysValPheLysHisLysAspLeuGlnGlnGln 319
DB 1006 TACGAGCTCGGAGAGAGACATCGACAAAGTCTTCAACACAGAGATCTGCGAGCAGCAG 1065
QY 320 LeuValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluGluArgHis 339
DB 1066 CTGGTGAGCGGCAAGCTCAGCAGCGGCGGAGAGTGTGAGGAGGAGGAGGAGGAGGAGGAG 1125
QY 340 GlnArgGluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeu 359

DB 1126 CAGCGGAGAGGAGACTTCTCTCGAGAGGAGCGCGTGAGTCCCGAGGAGATGTCCGAGCTG 1185
QY 360 MetLysGlnGlnGluThrHisLysLysGlnGluLeuAlaLeuThrThrGluLysPheGlu 379
DB 1186 ATGAAGCAACAGGAGACCCACTGAACAGCAGGCTTGCCTTATACACAGAGAGATTGAG 1245
QY 380 GluPheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMet 399
DB 1246 GAGTTCAGAGACACTCTTTCACAAAGCAGCGAGGTGTTCACACACATTCAACAGGAATG 1305
QY 400 GluLysMetThrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 419
DB 1306 GAAAGATGACAAAGAGATCAAGAGCTGAGAGAGAGACCCACCATGTACCGTTCCCGG 1365
QY 420 TrpGluSerSerAsnLysAlaLeuLeuGluMetAlaGluLysLysThrValaArgAspLys 439
DB 1366 TGGGAGACCAACAAAGGCGCTCTTGATGGCTGAGGAGAGAAACACTCCGGGACAAA 1425
QY 440 GluLeuGluGlyLeuGlnValLysLysLysLysLysLysLysLysLysLysLysLysLys 459
DB 1426 GAGCTGGAGGCGCTCCAGGTGAAATCCAGCGCTGGAGAGAGCTGCCCGGCGACTGCAG 1485
QY 460 ThrGluArgAsnAspLeuAsnLysArgValGluAspLeuSerAlaGlyGlyGlnGlySer 479
DB 1486 ACAGAGCGCATGACCTGAACAGAGGCTGAGGAGCTGAGTCCCGTGGCGGAGGCGGCG 1545
QY 480 LeuThrAspSerGlyProGluArgArgPro----- 489
DB 1546 GTCTCCGAGCGCTCTCGAGCGGAGCCAGAGCCCGCCACCTCCCAAGGAGCAGGCT 1605
QY 490 ---GluGlyProGlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyr 508
DB 1606 GTCGAGGCGCCCGGGCTCAAGTACCCACTCTCCAGGCGCCACAGACGCTTCTCTCTGC 1665
QY 508 ProGlyAlaProSerThrGluAlaSerGlyGluThrGlyProGlnGluProThrSerAla 528
DB 1666 GCAGGTGCACCCAGCAGAGGAGCATCAGGCGGAGGCGGCGGCGGCGGCGGCGGCGG 1725
QY 529 ArgAla 530
DB 1726 ACTGCC 1731

RESULT 15

US-10-671-242-14
; Sequence 14, Application US/10671242
; Publication No. US20040040049A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/671,242
; CURRENT FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 4722
; TYPE: DNA
; ORGANISM: Oryctolagus cuniculus

QY 529 ArgAla 530
 Db 1726 ACTGCC 1731
 RESULT 18
 US-10-616-187-14
 ; Sequence 14, Application US/10616187
 ; Publication No. US2004001368A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lees, Robert S.
 ; APPLICANT: Lees, Robert S.
 ; APPLICANT: Law, Simon W.
 ; APPLICANT: Arjona, Anibal A.
 ; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
 ; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
 ; TITLE OF INVENTION: ATHEROSCLEROSIS
 ; FILE REFERENCE: 10797-004001
 ; CURRENT APPLICATION NUMBER: US/10/616,187
 ; CURRENT FILING DATE: 2003-07-09
 ; PRIOR APPLICATION NUMBER: US/09/616,289
 ; PRIOR FILING DATE: 2000-07-14
 ; PRIOR APPLICATION NUMBER: US 09/517,849
 ; PRIOR FILING DATE: 2000-03-02
 ; PRIOR APPLICATION NUMBER: US 08/979,608
 ; PRIOR FILING DATE: 1997-11-26
 ; PRIOR APPLICATION NUMBER: US 60/031,930
 ; PRIOR FILING DATE: 1996-11-27
 ; PRIOR APPLICATION NUMBER: US 60/048,547
 ; PRIOR FILING DATE: 1997-06-03
 ; NUMBER OF SEQ ID NOS: 53
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 14
 ; LENGTH: 4722
 ; TYPE: DNA
 ; ORGANISM: Oryctolagus cuniculus
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (61)...(1731)
 US-10-616-187-14
 Alignment Scores:
 Pred. No.: 1,18e-190 Length: 4722
 Score: 2465.50 Matches: 492
 Percent Similarity: 92.99% Conservative: 12
 Best Local Similarity: 90.77% Mismatches: 25
 Query Match: 91.25% Indels: 13
 DB: 16 Gaps: 3
 US-10-023-523-8 (1-530) x US-10-616-187-14 (1-4722)
 QY 1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 20
 Db 109 AAAAGCAGCCCGGACAGCCGGAAGCAGGAGCGGAGCGCCAGCGGCGCGCGCGG 168
 QY 21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40
 Db 169 CCGGCCCCCGCCGAGAGCCGAGT---GCCAGCAGCCAGGCTCCCGGAGGCGCGAG 225
 QY 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuA:9ASpValSerGluGluLeuSer 60
 Db 226 GGGGCTCAAGCCAAACTCTCAGCCTGGGGCGCTCTGTGTGTCTCTGAGGAGCTGAGC 285
 QY 61 ArgGlnLeuGluAspIleuSerThrTyrCysValaspAsnAsnGlnGlyGlyProGly 80
 Db 286 CCGCAGTTGGAAGACATACTCTAGTACATACTGTGTGGACCAACAGGGGGCCCCGGGT 345
 QY 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
 Db 346 GAGGATGGGGTCCAGGGTCCAGCCCTGAACTGAGTGCAGAGAGTCTCGCGCTAT 405
 QY 101 ValAlaArgAsnGlyGluProGluPro----ThrProValValTyrGlyGluGluPro 119
 Db 1486 ACAGAGCGCAATGACCTGAACAGAGGGTTCAGGACCTGAGTCCGGTGGCCAGGGCCCC 1545
 Db 406 GTGCAAGCAATCGGAGCGGAGCGGCGACCCCAAGTGTCAATGCGGAGAGGAGACC 465
 QY 120 SerLysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAsp 139
 Db 466 TCCAAGCGCAGAGCCGGGCGGAGAGATCCGAGCAGCGATGAGTCCGAGACCGAGAC 525
 QY 140 HisArgArgProGlnGluLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeu 159
 Db 526 CACCGGAGGCCACAGAAAAGAGAGGCAAGGGTCTGGAAAGGAGATCACGCTGCTG 585
 QY 160 MetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLys 179
 Db 586 ATGACGACACTGAACACGCTGAGCAGCCAGAGAGAGAGCTGGCGCTCTGTGCAAGAG 645
 QY 180 TyrAlaGluLeuLeuGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLys 199
 Db 646 TATCGGGAAGTCTCGAGGAGCACCCGGAAGTCCGAGACAGATGAAGTGTGTCAGAG 705
 QY 200 LysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaVal 219
 Db 706 AAGCAGAGCCAGCTGTGCGAGGAGAGGACCACTGCTGGCGAGCAGCAGAGGCCATC 765
 QY 220 LeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeu 239
 Db 766 CTGGCCCGCAGCAGCTGCGAGAGCTGTGCGGAGAGCTGCGAGCGGACCAACCGCTGCTC 825
 QY 240 LysGluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSer 259
 Db 826 AAGAAAGAGGTGTGCGAGGAGCGCCGAGAGGAGGAGAGGAGGAGGAGGTGAGCTCA 885
 QY 260 HisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsn 279
 Db 886 CACTTCAGATGACGCTCAACGACATTCAGTGCAGTGGAGCAGCAGCAACAGGCGCAAC 945
 QY 280 SerLysLeuArgGlnGlnAsnMetGluLeuAlaGluArgLeuLysLysLeuLeuGluGln 299
 Db 946 TCCAAGCTGCGCCAGGAGAACATGAGCTGGCGGAGCGGCTCAAGAAGCTGATTGAGCAG 1005
 QY 300 TyrGluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGln 319
 Db 1006 TAGCAGCTGCGAGAAGAGCAGCATCGCAAGTCTTCAACACAGAGGATCTGAGCAGCAG 1065
 QY 320 LeuValAspAlaLysLeuGlnGlnAlaGlnMetLeuLysGluAlaGluArgHis 339
 Db 1066 CTGGTGAGCCCAAGCTCCAGCAGGCGCCAGGAGATGCTGAAGAGGAGCAGAGGCGGCAC 1125
 QY 340 GlnArgGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeu 359
 Db 1126 CAGCGGAGAGGAGACTTCTCTCAAGAGGAGCGCGTCCAGAGGATGTGCGAGCTG 1185
 QY 360 MetLysGlnGlnGluThrHisLeuLysGlnLeuAlaLeuTyrThrGluLysPheGlu 379
 Db 1186 ATGAAGCAACAGGAGACCCACCTGAAGAGCAGCTTGCCCTATACACAGAGATTGAG 1245
 QY 380 GluPheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMet 399
 Db 1246 GAGTTCCAGAACACTCTTTCAAAAGCAGCAGGTGTTCACCACTTCAACAGAGAAATG 1305
 QY 400 GluLysMetThrLysLysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArg 419
 Db 1306 GAAAAAGATGACAAAGAGAGATCAAGAGAGCTGGAGAAAGAGACCACTGATCCGTTCCCGG 1365
 QY 420 TrpGluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLys 439
 Db 1366 TGGAGAGCAGCAACCAAGGCCCTGCTTGAGTGGCTGAGGAGAAAAACACTCCGGGACAAA 1425
 QY 440 GluLeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGln 459
 Db 1426 GAGCTGGAAGCCTGCGAGGTGAAAAATCCAGCGGCTGAGAGAGCTGCGCGGCGACTGAG 1485
 QY 460 ThrGluArgAsnAspLeuAsnLysA:9ValGlnAspLeuSerAlaGlyGlyGlnGlySer 479
 Db 1486 ACAGAGCGCAATGACCTGAACAGAGGGTTCAGGACCTGAGTCCGGTGGCCAGGGCCCC 1545

QY 480 LeuThrAspSerGlyProGluArgArgPro----- 489
 DB : : : : :
 DB 1546 GTCTCCGACGCGCTCTTGAGCGGAGCCAGAGCCGCCACACCTCCCAAGAGAGAGGT 1605
 QY 490 ---GluGlyProGlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyr 508
 DB : : : : :
 DB 1606 GTCAGAGGCGCGGGGCTCAAGTACCACTCTCCAAGGCGCCACAGACGCTTCTCTGCTGC 1665
 QY 509 ProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAla 528
 DB : : : : :
 DB 1666 GCAGGTGCACCCAGCAGCAGAGGATCAGGCGACAGAGGCGCCCGAGGAGCCACCTGCG 1725
 QY 529 ArgAla 530
 DB : : : : :
 DB 1726 ACTGCC 1731

RESULT 19

US-10-115-831-34
 ; Sequence 34, Application US/10115831
 ; Publication No. US20030219743A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Ren, Feiyan
 ; APPLICANT: Dmanac, Radoje T.
 ; TITLE OF INVENTION: No. US20030219743A1el Nucleic Acids and

; TITLE OF INVENTION: Polypeptides
 ; FILE REFERENCE: 792CIP2ADIV
 ; CURRENT APPLICATION NUMBER: US/10/115,831
 ; CURRENT FILING DATE: 2002-04-02
 ; PRIOR APPLICATION NUMBER: 09/667,298
 ; PRIOR FILING DATE: 2000-09-22
 ; PRIOR APPLICATION NUMBER: 09/577,408
 ; PRIOR FILING DATE: 2000-05-18
 ; NUMBER OF SEQ ID NOS: 178
 ; SOFTWARE: pt_FL_genes Version 2.0
 ; SEQ ID NO 34
 ; LENGTH: 2356
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (85)..(1554)
 US-10-115-831-34

Alignment Scores:

Pred. No.: 1,78e-95 Length: 2356
 Score: 1290.50 Matches: 285
 Percent Similarity: 67.62% Conservative: 70
 Best Local Similarity: 54.29% Mismatches: 124
 Query Match: 47.76% Indels: 46
 DB: 16 Gaps: 9

US-10-023-523-8 (1-530) x US-10-115-831-34 (1-2356)

QY 13 GlyAlaGlnGluArgProSerGlnAlaAlaProAlaValGlnAlaGluGlyProGlySer 32
 DB 10 GCGCCGGAAGAG-----GCGACTGAGCGCGGAGCGGCGGAGCG 48
 QY 33 SerGlnAlaProArgLys-----ProGluGlyAlaGlnAlaArgThrAlaGlnSerGly 50
 DB 49 CGAGCGCGCGCGGCGAGAGTTTGAAATTGGCACATGGAAGAGCTGGAATTTGTGG 108
 QY 51 AlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyr 70
 DB 109 CTAGGGGTGAAGACAGATATGTTGTGTAACCTCTCAATCAATGATATCTTCAACATCAA 168
 QY 71 CysValAspAsnGlnGlyGlyProGlyGluAspGlyAlaGlnGlyProAlaGlu 90
 DB 169 -----GGCTCAAATTGTGGTGGCAGACAGTAACAAGCATTCTTGAAGAGGATGAAGGC 222

QY 91 ProGluAspAlaGluLysSerArgThrTyrValAlaArgAsnGlyGluProGluProThr 110
 DB : : : : :
 DB 223 AGTGACTTTATACAGAGACAGCAATTTGGTGAGC-----CCAGCA 264
 QY 111 ProValValTyrGlyGluLysGluProSerLysGlyAspProAsnThrGluGluIleArg 130
 DB : : : : :
 DB 265 TACTGCACGCAAGAAATCAAGAGAGAAATCCCTGGGGGA-----GAAAGCTCGA 312
 QY 131 GlnSerAspGluValGlyAspArgAspHisArgArgProGlnGluLysLysLysAlaLys 150
 DB : : : : :
 DB 313 ACAGATCCCTCATGGTCAGCAAGATTTCAGAGTGCAACAGGACCAAGAAAAAACT--- 369
 QY 151 GlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGlu 170
 DB : : : : :
 DB 370 ---TTAGGAAAGAAGTTTATTACTGATGCAAGCCCTAAACACCCCTTTCAACCCCGAG 426
 QY 171 GluLysLeuAlaAlaLeuCysLysLysTyrAlaGluLeuGluGluHisArgAsnSer 190
 DB : : : : :
 DB 427 GAGAGCTGGCAGCTCTCTGTAGAATAATGCTGATCTTCGGAGGAGAGGAGTGT 486
 QY 191 GlnLysGlnMetLysLeuLysLysGlnSerGlnLeuValGlnLysAspHis 210
 DB : : : : :
 DB 487 CAGAAGCAAAATGAAGATCTCTGCAGAAAGAAAGCAAGCCAGATTGTGAAGAGAAAGTTAC 546
 QY 211 LeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLysCysArg 230
 DB : : : : :
 DB 547 TTGCAGAGTGAACATAGCAGAGGCTATCTTGGCAGAGAGCAAGCTAGATCTTTTCGAGA 506
 QY 231 GluLeuGlnArgHisAsnArgSerLeuLysGluGlyValGlnArgAlaArgGluGlu 250
 DB : : : : :
 DB 607 GAATCTCAGCGTCACATTAAGACGTTTAAAGCGAGGAAATATGCGAGCAGCAGAGAGAA 666
 QY 251 GluLysLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeu 270
 DB : : : : :
 DB 667 GAGAGCAGCTAAAGAGCAACTGCACATTTCCAGATTACCTTAATGAATTCAGGC 726
 QY 271 GlnMetGlnHisAsnGluArgAsnSerLysLeuArgGlnGlnAsnMetGluLeuAla 290
 DB : : : : :
 DB 727 CAGCTGGAGCAGCATGACATCCACAAACGCGCAAACTCCGACAGGAAACATTTAGCTGGG 786
 QY 291 GluArgLeuLysLysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAspLysVal 310
 DB : : : : :
 DB 787 GAGAGCTAAAGAGCTCATCGAACAGTACGCTGAGGGAAGAGCAGCATTTGATAGGTG 846
 QY 311 PheLysHisLysAspLeuGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGlu 330
 DB : : : : :
 DB 847 TTCAAAACATAAAGAACTGCAACAGCAGCTCGTGGATGCCAAACTCGAGCAAAACGACAA 906
 QY 331 MetLeuLysGluAlaGluGluArgHisGlnArgGluLysAspPheLeuLysLysGluAla 350
 DB : : : : :
 DB 907 CTGATAAAGAGCTGATGAAACATCAGAGAGAGAGAGTTTATTATAAAGAGCG 966
 QY 351 ValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlnThrHisLeuLysGlnGln 370
 DB : : : : :
 DB 967 ACAGAAATCGAGGACAAATACGAAACAAATGAAACAGCAAGAGATCACTAAACAGCAG 1026
 QY 371 LeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGlu 390
 DB : : : : :
 DB 1027 CTCTCTCTTATATGATAGTTTGAAGATTCCAGACTACCATGCGCAAAAGCAATGAA 1086
 QY 391 ValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLysLeuGlu 410
 DB : : : : :
 DB 1087 CTGTTTACACCTTCAGACAGGAAATGCAAAAGATGCAAAAGAAATTAATAAACTGGAA 1146
 QY 411 LysGluThrThrMetTyrArgSerArgTTPGluSerSerAsnLysAlaLeuLeuGluMet 430
 DB : : : : :
 DB 1147 AAGAAACAAATAATTTGGCGTACCAATGGGAAACAAATAATAAGCACTTCTGCAATG 1206
 QY 431 AlaGluGluLysThrValArgAspLysGluLeuGluGlyLeuGlnValLysIleGlnArg 450
 DB : : : : :
 DB 1207 GCTGAAGAGAGAAACAGTCCGCTGATAAAGAGTAGTCAAGGCCCTTCAATAAACTGGAACG 1266
 QY 451 LeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgValGln 470

Db 1267 TTAGAGAGCTGTGACGGCTCTTCAACAGAAAGGATGAGTCAATGAGAGGTGGAA 1326
 QY |||||
 Db 471 -----AspLeuSer 473
 QY |||||
 Db 1327 GTCCTGAAAGAGCAGGTATCCATCAAGCGGCCCATCAAGCGCGAACACAGGGATTTAGCA 1386
 QY |||||
 Db 474 AlaGlyGlyGlnGlySerLeuThr-----AspSerGlyProGluArgArgProGluGly 491
 QY |||||
 Db 1387 ACACCTGTGATGCAGCCCTGCTACTGCCCTGATCTCACAGAGGCTGGAACATCTCTCG 1446
 QY |||||
 Db 492 Pro-----GlyAlaGlnAlaProSer-SerProArgValThrGluAlaProCysTy 508
 QY |||||
 Db 1447 AAAAGAGCCCTGGGAGCCACCTCGAGGCTGAGCCCAAGAGTCAAGAGCGCTGTGCAA 1506
 QY |||||
 Db 508 rProGlyAlaPro 512
 QY |||||
 Db 1507 AAGCCCCCGTCCA 1519

RESULT 20

US-09-960-253-172
 ; Sequence 172, Application US/09960253
 ; Patent No. US20020123619A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Benson, Darin R.
 ; APPLICANT: Mohamath, Raodoh
 ; APPLICANT: Lodes, Michael J.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.556
 ; CURRENT APPLICATION NUMBER: US/09/960.253
 ; CURRENT FILING DATE: 2001-09-20
 ; NUMBER OF SEQ ID NOS: 187
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 172
 ; LENGTH: 2045
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-960-253-172

Alignment Scores:
 Pred. No.: 4,2e-95 Length: 2045
 Score: 1285.00 Matches: 284
 Percent Similarity: 67.49% Conservative: 71
 Best Local Similarity: 53.99% Mismatches: 132
 Query Match: 47.56% Indels: 39
 DB: 9 Gaps: 8

US-10-023-523-8 (1-530) x US-09-960-253-172 (1-2045)

QY 12 GluGlyAlaGlnGluArgProSerGlnAlaAlaProAlaValGluAlaGluGlyProGly 31
 Db 70 GAGGACGCGCGGGAAGAGGCGCGCGCGGCGGAGAGCGGAGGCGGAGCGGCGGCGGAG 129
 QY 32 SerSerGlnAlaProArgGlyS-----ProGluGlyAlaGlnAlaArgThrAlaGlnSer 49
 Db 130 CGGCGACCGCGCGCGGAGAGGCTTCAATGGCAATGGCAATGGCAATGGCAATGGCAAT 189
 QY 50 GlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGluAspLeuSerThr 69
 Db 190 GGGCTAGGGGTCAAGACAGATATCTGTGAACCTCAATCAATCAATCAATCAATCAATCA 249
 QY 70 TyrCysValAspAsnAsnGlnGlyProGlyGluAspGlyAlaGlnGlyProAla 89
 Db 250 CNA-----GGCTCAATTTGGTGGCAAGTCAAGCAAGCAATTCATGGAGAGGATGAA 303
 QY 90 GluProGluAspAlaGluLysSerArgThrTyValAlaArgAsnGlyGluProGluPro 109
 Db 304 GCGAGTGTCTTTATAACAGAGAACAGGAATTTGGTGAGC-----CCA 345
 QY 110 ThrProValValTyGlyGluLysGluProSerSerGlyAspProAsnThrGluGlu 129
 Db 346 GCATACTGCACGAAGATCAAGAGAGGAAATCCCTGGGGA-----GAAGCT 393

QY 130 ArgGlnSerAspGluValGlyAspArgAspHisArgArgProGlnGluLysLysLysAla 149
 Db 394 CGAACAGATCCCTCGATGCTCAGCAAGATTAGAGTGCACAGGAAACAAGAAAAACT 453
 QY 150 LysGlyLeuGlyLysGluLeuThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrPro 169
 Db 454 -----TTAGGAAAGAGAGTTTATTACTGATGCAAGCCCTAAACACCTTTTCAACCCCA 507
 QY 170 GluLysLeuAlaAlaLeuCysLysLysTyValAlaGluLeuLeuGluLysHisArgSer 189
 Db 508 GAGGAGAGCTGCGAGCTCTCTGTAAGAAATATGCTGATCTTCTGGAGAGGAGGAGGT 567
 QY 190 SerGlnLysGlnMetLysLeuGlnLysLysGlnSerGlnLeuValGlnGlnLysAsp 209
 Db 568 GTTCAGAGCAATGAAGATCTCTGCAGAGAGAGCAAGCCAGATTGTGAAGAGAGAGTT 627
 QY 210 HisLeuArgGlyGluHisSerLysAlaValLeuAlaAspSerLysLeuGluSerLeuCys 229
 Db 628 CACTTGCAGAGTGAACATGCAAGGCTATCTTGGCAAGAGCAAGCTAGATTCTCTTTTGC 687
 QY 230 ArgGluLeuGlnArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGlu 249
 Db 688 AGAGAACTTCAGGCTCACAAATAGACGTTAAAGAGGAGGAGGAGGAGGAGGAGGAG 747
 QY 250 GluLysLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspLeuGln 269
 Db 748 GAGAGAGAGCGTAAAGAGCAACTGCACATTTCCAGATTACCTTAGATGAGAAATTCAA 807
 QY 270 LeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGlnGlnMetGluLeu 289
 Db 808 GCCCAGCTGGAGCAGCATGATCCACACAGCCCAACTCCGACAGGAGGAGGAGGAGGAG 867
 QY 290 AlaGluArgLeuLysLysLeuLeuGlnGlnTyGluLeuArgGluGluHisLysLys 309
 Db 868 GGGGAGAGCTAAAGAGCTCATCGAACAGTACGCACTGAGGAGAGAGGAGGAGGAGGAG 927
 QY 310 ValPheLysHisLysAspLeuGlnGlnGlnLeuValAlaLysLeuGlnGlnAlaGln 329
 Db 928 GTGTTCAACGTAAGGAACTGCAACAGCAGCTCGTGGATGCCAAACTCGACGAGGAGGAG 987
 QY 330 GluMetLeuLysGluAlaGluArgHisGlnArgGluLysAspPheLeuLeuLysGlu 349
 Db 988 CAACTGATAAAGAGCTGATGAAACATCAGAGAGAGAGAGGAGGAGGAGGAGGAGGAG 1047
 QY 350 AlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlnThrHisLysLys 369
 Db 1048 GCGACAGATCGAGGCGCAAAATACGAAACAAATGAAACACAGCAGGAGGAGGAGGAG 1107
 QY 370 GluLeuAlaLeuTyThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSer 389
 Db 1108 CAGCTTCTCTTTATATGATAGTTTGAAGATTCCAGACTCCATCCATGGCAAGAGCAAT 1167
 QY 390 GluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLys 409
 Db 1168 GAACGTGTTTCAACCTTCAGACAGGAAATGGAAGAGATGACAAAGAGGAGGAGGAGGAG 1227
 QY 410 GluLysGluThrThrMetTyArgSerArgTyrGluSerSerAsnLysAlaLeuLeuGlu 429
 Db 1228 GAAAAAGAAACAATAATTTGGCGGTACCAATGGGAAACAAATATTAAGACATCTTCTGCA 1287
 QY 430 MetAlaGluLysThrValArgAspLysGluLeuGluGlyLeuValLysLysLysLys 449
 Db 1288 ATGGCTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1347
 QY 450 ArgLeuLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgVal 469
 Db 1348 CGGTAGAGAGCTGTGCGGGCTCTTCAACAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1407
 QY 470 Gln-----AspLeu 472
 Db 1408 GAAGTCTCTGAAGAGGAGGAGGATCTCCATCAAGAGCGGCCCATCAAGAGCGGCGGAGGAG 1467

QY 401 ----- 401
Db 15412 GAATTACAGCGGGTCCACCATGCTGGCTAATTTTCTTTCTTTTGTGATTT 15471
QY 401 ----- 401
Db 15472 TTAGTAGAGAGGGGTTTACCATGTTGGCCAGGCTGGTCTCGAACTCTTGGCTCAAGT 15531
QY 401 ----- 401
Db 15532 GATCTGCCCGCTTGGCTCCCAAAGTCTGGGATTATAGGCGTGAGCCACCATGCTGG 15591
QY 401 ----- 401
Db 15592 CCGCTTACCATTCCTGTTTATGTTGGTGTGACACCTCTGACTCTCTGCTGGTGTGAGTGGC 15651
QY 401 ----- 401
Db 15652 ACAGAGGCGATTGACTGCTATCTGTAATGCTTGGCGCTTGGGATCAATCATTTCCCCACC 15711
QY 401 ----- 401
Db 15712 TTGAGACACAGGTGACGTCCCACTTTGGAGACACAGACCTTGGAGAGCCAGCTCTGA 15771
QY 402 ----- MetThrLysLysLysLysLysLys 411
Db 15772 CCATTCTCTGCTGTCTCATATAACCTAGATGACTTAAGAGATCAGAGCTGGAGAA 15831
QY 412 GluThrThrMetTyrArgSerArgTTPGluSerSerAsnLysAlaLeuLeuMetAla 431
Db 15832 GAAACCAACCATGTACCGTCCCGTGGGAGAGCAGCAACAGCGCTCTTGGATGGCT 15891
QY 432 GluGlu ----- 433
Db 15892 GAGGAGTGGCTGTCTGTGATCTGCAGCCAGGCTGGGGTGTGCTACTTAGCCCATATCA 15951
QY 433 ----- 433
Db 15952 GGCCTTCTCTGTATGTTCTACCCATCAGTGACACAGCTAGCATGAGGTAGGTGAGAT 16011
QY 433 ----- 433
Db 16012 TTGCACAAATGTCCAAGTCCAAAGTTAATGCTTCTCTCCCATGGGAGGTGAGC 16071
QY 433 ----- 433
Db 16072 CCAGTGTAGTCTCCAGTGGAGTGAAGGAGCAATGGAAGAGGATAAAGAGCA 16131
QY 433 ----- 433
Db 16132 GAAAAAAGCGTCCAGTGATGCTGCTGTTTACATGTAAAGCAGCCAGGTAGTTGT 16191
QY 433 ----- 433
Db 16192 GATTTCAGCTTGTATGTAGAGAAAGGAACTAACGATGGAGCAGCAACTGCAAGCCA 16251
QY 433 ----- 433
Db 16252 GACCTTGTGAAAGTTTTTGGTTTTTTTGTCTTTTGTCTGTGTAATTTTTTAGGTA 16311
QY 433 ----- 433
Db 16312 CGTTGTTTCAATTAACCTTCTCTGAGCTCTGAGATGATTAAGTCTCTGTTTATAG 16371
QY 433 ----- 433
Db 16372 ATGAGACAGGCTCAAAAGTCAAGTCTTTGCCAGGTCACTGGTAGATAAATGAGGAA 16431
QY 433 ----- 433
Db 16432 TACGTTATCCAAAGCGTGCCCTTTTCTGCACCATGCTGCCCCCACTGACAGCCTAGT 16491

QY 433 ----- 433
Db 16492 CATGCTTCAACTAGGACTGTTTCTAAAGGGGCCAGCTTTGGACTCGTCTGCTCTCA 16551
QY 433 ----- 433
Db 16552 GCTTGTAAAGTGTTCGCCCAAGTGGTGTATGTAAGTGGAGGTTGATGGGCACGG 16611
QY 434 ----- LysThrValArgAspLysGluLeuGluGlyLeu 444
Db 16612 CACTGAAGTCTCAATTTTCTTCCCTAGAAAAACAGTCCGGGATAAAGAACTGGAGGCGCTG 16671
QY 445 GlnValLysIleGlnArgLeuLysLeuCysArgAlaLeuGlnThrGluArgAsnAsp 464
Db 16672 CAGGTAAATCAACAGCTGGAGAGCTGTGCCGGGCACTGCAGACAGAGCGCATGAC 16731
QY 465 LeuAsnLysArgValGlnAspLeuSerAlaGlyGlnGlySerLeuThrAspSerGly 484
Db 16732 CTGAACAAGAGGTACAGACCTGAGTGTGTCGCCAGGCTCCCTCACTGACAGTGGC 16791
QY 485 ProGluArgGpProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGlu 504
Db 16792 CTTGAGAGAGGCCAGAGGGCTCGGGCTCAAGCACCAGCTCCCCAGGCTCACAGAA 16851
QY 505 AlaProCysTyrProGlyValaProSerThrGluAlaSerGlyGlnThrGlyProGlnGlu 524
Db 16852 GCGCTTGTCTACCCAGGAGCACCAGACAGAGCATCAGGCCAGACTGGGCTCAAGAG 16911
QY 525 ProThrSerAlaArgAla 530
Db 16912 CCCACCTCCGCCAGGGCC 16929
RESULT 23
US-10-671-242-51
; Sequence 51, Application US/10671242
; Publication No. US20040040049A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arizona, Arribal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/671,242
; PRIOR FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 22255
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-671-242-51
Alignment Scores:
Pred. No.: 4,71e-48 Length: 22255
Score: 721.50 Matches: 224
Percent Similarity: 23.78% Conservative: 1
Best Local Similarity: 23.68% Mismatches: 1
Query Match: 26.70% Indels: 721
DB: 13 Gaps: 4


```
QY 433 ----- 433
Db 16252 GACCTTGCTGAAGTGTGTTGGGTTTCTTTTGTCTTTTGTGCTGAATGTTTATAGGTA 16311
QY 433 ----- 433
Db 16312 CGTGTGTTCAATGAACCTTCTCTTGAGCTCTGAGGATGTTATAGTAGTCTCTGTTTATAG 16371
QY 433 ----- 433
Db 16372 ATGAGCAGGCTCAAAAGTCAAGTCTTTGTCACAGCTCAGGTGATAGATAAATGAGAGAA 16431
QY 433 ----- 433
Db 16432 TACGTTATCTCAAGCGTGCCCTTTCTGACCATGTGCCCCACCTGACAGCCTAGT 16491
QY 433 ----- 433
Db 16492 CATGCTTCAACTAGGACTGTTTCTTAAGGGGCCAGCTTTGGACTCGTCTCTCA 16551
QY 433 ----- 433
Db 16552 GCCTGTGTTAAAGTGTTCGCCCAAGTGTGTAAGTGGAGGTTGATGGGCACGG 16611
QY 434 -----
Db 16612 CACTGAAGGTCTCATTTCTTTCCCTAGAAACAGTCCGGGATAAGAACTGGAGGCGCTG 16671
QY 445 GlnValIysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAsp 464
Db 16672 CAGGTAAATCCAAACGCTGGAGAGCTGTGCGGGCACTGCAGACAGAGCGCAATGAC 16731
QY 465 LeuAsnLysArgValGlnAspLeuSerAlaGlyGlnGlySerLeuThrAspSerGly 484
Db 16732 CTGAACAAGAGGTACAGGACCTGAGTGTGTGGCCAGGGCTCCCTCACTCAGAGTGGC 16791
QY 485 ProGluArgArgProGluGlyProGlyAlaGlnAlaProSerProArgValThrGlu 504
Db 16792 CCTGAGAGGAGGCCAGAGGGGCTCGGGCTCAAGACCCAGCTCCCGGAGGTACAGAA 16851
QY 505 AlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGlu 524
Db 16852 GCGCTTGTCTACCCAGGAGACCCAGACAGACAGATCAGGCGAGCTGGGCTCAAGAG 16911
QY 525 ProThrSerAlaArgAla 530
Db 16912 CCCACCTCCCGCAGGGCC 16929

RESULT 24
US-10-023-529-51
; Sequence 51, Application US/10023529
; Publication No. US20020129388A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/023,529
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
```

```
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 22255
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-023-529-51

Alignment Scores:
Pred. No.: 4,71e-48 Length: 22255
Score: 721.50 Matches: 224
Percent Similarity: 23.78% Conservative: 1
Best Local Similarity: 23.68% Mismatches: 1
Query Match: 26.70% Indels: 721
DB: 14 Gaps: 4

US-10-023-523-8 (1-530) x US-10-023-529-51 (1-22255)

QY 305 GluHisIleAspLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLys 324
Db 14093 CAGCATATCGACAAAGTCTTCAACACAAAGGACCTACACAGCAGCTGGTGGATGCCAAG 14152
QY 325 LeuGlnGlnAlaGlnGlnMetLeuLysGluAlaGluGluArgHisGlnArgGluLysAsp 344
Db 14153 CTCACGAGGCCCGCAGAGATGCTTAAGAGGAGGACAGAGCGGCACCGCCGGGAGAGAT 14212
QY 345 Phe----- 345
Db 14213 TTTGTGAGGCTCAGGCCCGCAGGCTTGGGGTGGGGTGTGGGAGAGACAGGCTGGGCTCT 14272
QY 345 ----- 345
Db 14273 GGCTCAGCTCATAGCGGGTGTATATGGGAGAAAGTCTGGCCACAGCAGACAGATTCCTT 14332
QY 345 ----- 345
Db 14333 GAGTACCAGTCTCAGAGCAGGAAGCCTCAGTGGTCTGTGGCTTGTGGCTAAAAACCAA 14392
QY 346 -----LeuLeuLys 348
Db 14393 CATAGCCCTGGGGGCTTCTGACAGGATCTGGGTTCTGTCTTGGAAATAGTCTCTGAAA 14452
QY 349 GluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLeuLys 368
Db 14453 GAGCGAGTAGTCCCGAGGAGTGTGTAGCTGTATGAGCAGCAGCAGCAGCAGCAGCAG 14512
QY 369 Glu----- 369
Db 14513 CAACAGTGAGACATATAAACCCTGTCCTTCAAGTTTCCCTCACTGGGCCCCCAT 14572
QY 369 ----- 369
Db 14573 CCTGGGGTAGTGAATGGACCCCTCATTTAGGACTGGCTGTGTCTCTGCTGTATGAC 14632
QY 369 ----- 369
Db 14633 GCCTTGGTTAGCTTAGTGGGCTCAGAGGACTTCATTGTAGCTCAGAAATGATTGCT 14692
QY 369 ----- 369
Db 14693 TTTGAGAGTAGGAAACAGAGAGTTTGAATAATCAACATAAGGCAAAATAAAGTCACC 14752
QY 369 ----- 369
Db 14753 CTAAGTCTCTACTTTTCAGGCTTAGCATTTTGGATTATATCTTCCAAATATATAGCTT 14812
QY 369 ----- 369
Db 14813 TGCTTTGTTTAAAGGAAAAATAGTATCTCAATAGAATTACTTGGTCAGAGAGTCAAGGACG 14872
QY 369 ----- 369
Db 14873 GGTCTGAGTGTGTGACAGAGTGGCTCCACAGAGAAACCCAGTCTTATCTGTGGGCTGCT 14932
```


QY 402 -----MetThrLysLysLysLysLysLeuGluLys 411
Db 15772 CCATTTCCTTCGTCTGTGCATCAATCACTAGATGACTAAGAGATCAAGAAGCTGGAGAA 15831
QY 412 GluThrThrMetTyrArgSerArgTTPGluSerSerAsnLysAlaLeuLeuGluMetAla 431
Db 15832 GAACACACATGTACCGGTCCCGTGGGAGAGACAGCAACAGGCCCTGCTTGAGATGGCT 15891
QY 432 GluGlu----- 433
Db 15892 GAGGAGGTGGGCTGCTGTGTGATCTGCAGCCAGGCTGGGGGTGTGCACTTAGCCCATATCA 15951
QY 433 ----- 433
Db 15952 GGCCCTTCCTGTATGTCTTACCATCATCAGTGCACAGCTAGCATGAGGTAGAGGTGAGAT 16011
QY 433 ----- 433
Db 16012 TTGCACACATGTCCAGTCCAAAGTTAATGCTGTCTCTCCCATGGAGGTGGTGAGC 16071
QY 433 ----- 433
Db 16072 CCAGTGGTAGTCTCCAGTGGGAGTGAAGGAGCAATGAAGAAAGAAATAAAGAGCA 16131
QY 433 ----- 433
Db 16132 GAAAAAAGCGGTGCCAGTGTGCTGCTGCTGTTTACATGTAAAGCAGCCAGGTAGTTGT 16191
QY 433 ----- 433
Db 16192 GATTTACAGCTGTGAATGTAGAAAGAAAGAACTAACGATGGAGCAGCACTCAAGCCA 16251
QY 433 ----- 433
Db 16252 GACCTTGCTGAAGTTTTTGGGTTTTTTTGTCTTTTGTCTGCTGAATGTTTTTAGGTA 16311
QY 433 ----- 433
Db 16312 CGTTGTTCAATGAACCTTCTCTTGAGCTCTGAGGATGTTATTAGTACTCTGTTTTTAG 16371
QY 433 ----- 433
Db 16372 ATGACACAGGCTCAAAAGTCAAGTCTTGTCCAAAGTCACTGCTGATAGTAATATGGAGAA 16431
QY 433 ----- 433
Db 16432 TACGTTATCTCAAGCGGTGCCCTTTTCTGCACCATGTGCCCCACCTGACAGCCTAGT 16491
QY 433 ----- 433
Db 16492 CATGCTCAACTAGGACTGTTCTTAAAGGGGGCCAGCTTTGGACTCGGTCTGCTCTCA 16551
QY 433 ----- 433
Db 16552 GCCTTTGTTAAAGTGTTCGCCCAAGTGTGTGATGTAAGTGGGAGGTGATGGGGCACGG 16611
QY 434 ----- 444
Db 16612 CACTGAGGTCTCATTTCTTCCCTAGAAAACAGTCCGGGATAAAGAACTGGAGGGCTG 16671
QY 445 GlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAsp 464
Db 16672 CAGTAAAAATCCAAACGGCTGGAGAAGCTGTGGCGGCACTGACAGACAGAGGCGCATGAC 16731
QY 465 LeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSerGly 484
Db 16732 CTGACACAGAGGTACAGGACCTGAGTGTCTGTGGCCAGGGCTCTCTCACTGACAGTGGC 16791
QY 485 ProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGlu 504
Db 16792 CCTGAGAGGAGGCCAGAGGGGCTGGGGCTCAAGCACCCAGCTCCCGGCTCAGAA 16851
QY 505 AlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGluGlu 524

Db 16852 GCGCTTGCTTACCAGGAGCACCCAGGACACAGAAGCATCAGGCAGACTGGGCTCAAGAG 16911
QY 525 ProThrSerAlaArgAla 530
Db 16912 CCCACCTCCGCCAGGGCC 16929
RESULT 26
US-10-616-187-51
; Sequence 51, Application US/10616187
; Publication No. US20040013668A1
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/10/616,187
; CURRENT FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 22255
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-616-187-51
Alignment Scores:
Pred. No.: 4,71e-48 Length: 22255
Score: 721.50 Matches: 224
Percent Similarity: 23.78% Conservative: 1
Best Local Similarity: 23.68% Mismatches: 1
Query Match: 26.70% Indels: 721
DB: 16 Gaps: 4
US-10-023-523-8 (1-530) x US-10-616-187-51 (1-22255)
QY 305 GluHisIleAspLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLys 324
Db 14093 CAGCATATCCAAAGTCTTCAACACAGGACCTACACACAGCTGTGTGATGCCAAG 14152
QY 325 LeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluGluArgHisGlnArgGluLysAsp 344
Db 14153 CTCAGCAGGCCCCAGGAGATGCTTAAAGGAGGAGAGAGAGCGCCACAGCGGAGAGGAT 14212
QY 345 Phe----- 345
Db 14213 TTTGTGAGGCTCAGGCCCCAGGCTTGGGGTGTGGGAGGAGACAGGCTGGGCTCT 14272
QY 345 ----- 345
Db 14273 GGCTCAGCTCATAGCCGGGTTATATGGGAGAGTCTGGCCAGACGACAGATTCCTT 14332
QY 345 ----- 345
Db 14333 GAGTACCAAGTCTGAGACGAGGAAGCCCTCAGTGGGTCTGTGTGCTTGGCTAAACACAAA 14392
QY 346 -----LeuLeuLys 348
Db 14393 CATAGCCCTGGGGGCTTCTGTACAGGATCTGGGTTCTGTCTTGGAAATAGCTCTCTGAAA 14452

Qy	401	-----	401
Db	15532	GATCTGCGCCTTGGCCTCCAAAGTGTGGGATTATAGGCGTGAGCCACCATCGCTGG	15591
Qy	401	-----	401
Db	15592	CCCTTACCATTCTTGTATTATGGTGGTGACACCTCTGACTTCTGGTGTGAGTGCG	15651
Qy	401	-----	401
Db	15652	ACAGAGGCATTGACTGCATCCTGTAATGCCTTGGCCTTGGGATCAATCATTCGCCACC	15711
Qy	401	-----	401
Db	15712	TTGGAGACACAGGTGAGTCCCACTTGGAGACACAGACCTTGGAGAGGCCAGCTCTGA	15771
Qy	402	-----MetThrLysLysIleLysLysLeuGlulys	411
Db	15772	CCATTTCTTCGTCTGTACATAACCTAGATGACTAAGAAGATCAAGAAGCTGAGAAA	15831
Qy	412	GlulThrThrMetTyrArgSerArgTrpGluSerSerAsnLysAlaLeuLeuGluMetAla	431
Db	15832	GAACACCACTGTACCGGTCCCGGTGGGAGACAGCAAACAAAGCCCTCTTTGAGATGGCT	15891
Qy	432	GlulGlu-----	433
Db	15892	GAGGAGTGGCTGTCTGTGATCTGCAGCCAGGGTGGGGTGTGCATTCAGCGCATATCA	15951
Qy	433	-----	433
Db	15952	GGCCCTTTCTGTATGTTCTTAACCATCAGTGACACAGCTAGCATGAGGTAGAGGTAGAT	16011
Qy	433	-----	433
Db	16012	TTGCACACAATGTCCAAGTCCAAAGTTAATGCTGTTCTCTCCCATGGAGGTGTGAGC	16071
Qy	433	-----	433
Db	16072	CCAGTGTAGGTCTCCAGTGGGAGTGAGGGAGCAAATGGAAGAAAGGAATAAAAGAGCA	16131
Qy	433	-----	433
Db	16132	GAATAAAACGGGTGCAGTGATGTGCTGTTTACATATAAGACGCCAGCTGATTTGT	16191
Qy	433	-----	433
Db	16192	GATTTACAGCTTGTAAATGTAGAGAAAGGAACTAACGATGGAGCAGCAACTGCAGGCCA	16251
Qy	433	-----	433
Db	16252	GACCTTGCTGAAAGTTTTTGGGTTTTTTTTTGTCTTTTGTGCTGAATGTTTTTAGTA	16311
Qy	433	-----	433
Db	16312	CGTGTTCATTGACCTTCTCTTGAGCTCTGAGGATGGTATTAGTAGTCTGTTTATTAG	16371
Qy	433	-----	433
Db	16372	ATGAGACAGCTCAAAAGTCAAAGTCTTTGCCAAGTCACTGCTGAGTATAAATGGAGAA	16431
Qy	433	-----	433
Db	16432	TAGCTTATCCAGCCGTGGCCCTTTTCTGCACCACTGCTGCCCACTGACGCCCTAGT	16491
Qy	433	-----	433
Db	16492	CATGGCTTCAACTAGGACTGTTTCTTAAAGGGGGCCACTTTGGACTCGGTCTGCTCTCA	16551
Qy	433	-----	433
Db	16552	GCCTTGTAAAGTGTTCGCCCAAGTGGTATGATGGAGGTTGATGGGGCAGCG	16611
Qy	434	-----lvstThrValArgAspLysGluLeuGlulvLeu	444

Db	22	GGTGGCGGCGCTCTAGAACTAGTGGATCCCCC---GGGCTGCAGGAATT-CGGCACGAG	77
Qy	49	SerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGluAspIleLeuSer	68
Db	78	GCTGGAGCCCTCCGTGATGTCCTGTAGAGATTGAGCCGCAGCTGGAGACATCCTCAGT	137
Qy	69	ThrTyrCysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAlaGlnGlyGluPro	88
Db	138	ACATACTGCGTGACAAACAGTCAAGGGGGCCAGGTGAGGATGTGGCACAGGGTGAACCT	197
Qy	89	AlaGluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsnGlyGluProGlu	108
Db	198	GCTGAACCCGAGATGCGAGAAAGTCCCGGACCTATGCCTCAAGAAATGGGGAGCCCTGAG	257
Qy	109	Pro---ThrProValValTyrGlyGluLysGluProSerLysGlyAspProAsnThrGlu	127
Db	258	CCAGAGACTCCAGTAGTCATGTTGAGAAGAAATCTCCAAGGGGGAGCCGGGCCCGGAC	317
Qy	128	GluIleArgGlnSerAspGluValGlyAspArgAspHisArgArgProGlnGluLysLys	147
Db	318	GAGATCCGGACCCAGTGATGAAGTCGTAGACCGAGACCCACCGAAGGCCACAGGANAAGAA	377
Qy	148	LysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThr	165
Db	378	AAAGCCAAAGGGTCTGGGAAGAGANATACCC-----TCCGCCACTTTAACTAC	425

US-09-918-995-18399/c
; Sequence 18899, Application US/09918995
; Publication No. US20030073623A1

GENERAL : APPLICATION

```

; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS GENA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FASTSEQ FOR WINDOWS Version 3.0
; SEQ ID NO 18899
; LENGTH: 334
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-18899

Alignment Scores:
Pred. No.: 2.77e-29 Length: 334
Score: 464.00 Matches: 87
Percent Similarity: 91.89% Conservative: 15
Best Local Similarity: 78.38% Mismatches: 9
Query Match: 17.17% Indels: 0
DB: 10 Gaps: 0

US-10-023-523-8 (1-530) x US-09-918-995-18899 (1-334)

```

30

274	AGAGAACTTCAGCGTCACAAATAAGACGCTTAAGAGAGGAAATATGACGAGCAGCACGAGAG	215
250	
Qy	250 GluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspLeuGln	269
Db	
Qy	214 GAAGAAGAACACGCTAAAGAGCAGCAACTGCACATTCGAGATTACCTTAATGAATTCGAA	155
Db	
Qy	270 LeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGluLeu	289
Db	
Qy	154 GCCCAGCTGGGAGCAGCATGCATCCACAAGCCCAAACTCCGACGAGGAAACATTTGACGCTG	95
Db	

Qy	326	GlnGlnAlaGlnGluMetLeuLysGlnAlaGluGluArgHisGlnArgGluLysAspPhe	345
Db	493	CAACAAACATCAAGAG-----AAGCGTCTCAGGAACATACACAATGCAG-----	537
Qy	346	LeuLeuLysGluAlaValGluSerClnArgMetCysGluLeuMetLysGlnGlnGluThr	365
Db	538	-----TTGATGCTTGAAACAAGATTCTCAGCTTATG---ACTACTCGAAG	579
Qy	366	HisLeuLysGlnGlnLeuAlaLeuTyThrGluLysPheGluPheGlnAsnThrLeu	385
Db	580	AACTGCGGTTCACACTAGCTTCTGACGGGAAAGATTTCAGCACTTTCAGGATGCTTG	639
Qy	386	SerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLys	405
Db	640	TCAAAAAGCAATGAAGTCTTTGAAACCTTACAAAGCAGGAGATGAAAGATGATTTCCAGTG	699
Qy	406	IleLysLysLeuGluLysGluThrThrMetTyArgSerArgTyrGluSerSerAsnLys	425
Db	700	ATAAAGATCTTARGAAGGAGACGAATTTCTGAAGGGAAAATGTGAGAACTCAGATATT	759
Qy	426	AlaLeuLeuGluMetAlaGluLysThrValArgAspLysGluLeuGluGlnLeuGln	445
Db	760	GCTATTGTGAAGCTCATTTGAAGACGCTGAGCTAACAAAGAAAGCAAAATAGAGAAATTGAAA	819
Qy	446	ValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeu	465
Db	820	AATCAAGGAGAGAGCTCGATCCCTGTGTGCAACACTACAGCAGAAAGGAAA-----	873
Qy	466	AsnLysA-gValGlnAspLeuSerAlaGlyGlnGlySerLeuThrAspSerGlyPro	485
Db	873	-----	873
Qy	486	GluArgArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGlu	504
Db	874	-----CAAGGCCCTTCGCCAGTATTCAGATGCCCTTCTAGCCCAAGAA	918

RESULT 34
 US-09-764-869-467
 ; Sequence 467, Application US/09764869
 ; Patent No. US20020061521A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PC007
 ; CURRENT APPLICATION NUMBER: US/09/764,869
 ; CURRENT FILING DATE: 2001-01-17
 ; Prior application data removed - refer to PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 2442
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 467
 ; LENGTH: 405
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (377)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (398)
 ; OTHER INFORMATION: n equals a,t,g, or c
 US-09-764-869-467

Alignment Scores:		
Pred. No.:	6,48e-20	Length: 405
Score:	350.00	Matches: 83
Percent Similarity:	79.09%	Conservative: 4
Best local Similarity:	75.45%	Mismatches: 17
Query Match:	12.95%	Indels: 6
DB:	9	Gaps: 2

US-10-023-523-8 (1-530) x US-09-764-869-467 (1-405)

QY 57 GluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnGln 76
Db 75 AAGAAGCGGGCTGCCAACTGGAAGACATCTAGCACATCTGTGTGGACATAACACG 134
QY 77 GlyGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLys 96
Db 135 GGGGGCCCCGGCGAGATGGGGACAGGTGAGCGCGCTGACCCCGAAGATGCAGAGAG 194
QY 97 SerArgThrTyrValAlaArgAsnGlyGluPro-GluProThrProVal---ValTyrGln 115
Db 195 TCCCGGACCTATGTGGCAAGATGGGAGCCTTGAACCACTTCCAKTAGTTCATTTGG 254
QY 115 yGluLysGluProSerLysGlyAsp-ProAsnThrGluGluIle-ArgGlnSer-AspGln 134
Db 255 AGAAGAGGACCTTCCAAAGGGGATTCCAAACACAGAGAGATTCGGCAGAGTTGACGA 314
QY 134 uValGlyAspArgAspHis---ArgArgProGlnGluLysLysLysLysLysLysLys 153
Db 315 AGTTCGGAGCAAGGAACCATTCGAAGGCCACAGGAGAGAAAGAAAGCCCAAGGTTTGGG 374
QY 153 yLysGluIleThrLeuLeuMet 160
Db 375 GANGAGATCACGTTTGTCTGATT 396
RESULT 35
US-10-091-504-467
; Sequence 467, Application US/10091504
; Publication No. US20030059908A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C1
; CURRENT APPLICATION NUMBER: US/10/091,504
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2442
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 467
; LENGTH: 405
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (377)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (398)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-091-504-467
Alignment Scores:
Pred. No.: 6,48e-20 Length: 405
Score: 350.00 Matches: 83
Percent Similarity: 79.09% Conservative: 4
Best Local Similarity: 75.45% Mismatches: 17
Query Match: 12.95% Indels: 6
DB: 15 Gaps: 2
US-10-023-523-8 (1-530) x US-10-091-504-467 (1-405)
QY 57 GluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnGln 76
Db 75 AAGAAGCGGGCTGCCAACTGGAAGACATCTAGCACATCTGTGTGGACATAACACG 134
QY 77 GlyGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLys 96
Db 135 GGGGGCCCCGGCGAGATGGGGACAGGTGAGCGCGCTGACCCCGAAGATGCAGAGAG 194
QY 97 SerArgThrTyrValAlaArgAsnGlyGluPro-GluProThrProVal---ValTyrGln 115
Db 195 TCCCGGACCTATGTGGCAAGATGGGAGCCTTGAACCACTTCCAKTAGTTCATTTGG 254
QY 115 yLysGluIleThrLeuLeuMet 160
Db 375 GANGAGATCACGTTTGTCTGATT 396

Db 255 AGAAGAGGAACTCCCAAGGGGATTCCAAACACAGAGAGATTCGGCAGAGTTGACGA 314
QY 134 uValGlyAspArgAspHis---ArgArgProGlnGluLysLysLysLysLysLysLys 153
Db 315 AGTTCGGAGCAAGGAACCATTCGAAGGCCACAGGAGAGAAAGAAAGCCCAAGGTTTGGG 374
QY 153 yLysGluIleThrLeuLeuMet 160
Db 375 GANGAGATCACGTTTGTCTGATT 396
RESULT 36
US-10-227-577-467
; Sequence 467, Application US/10227577
; Publication No. US20040005575A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C2
; CURRENT APPLICATION NUMBER: US/10/227,577
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 10/091,504
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 09/764,869
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 467
; LENGTH: 405
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (377)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (398)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-227-577-467
Alignment Scores:
Pred. No.: 6,48e-20 Length: 405
Score: 350.00 Matches: 83
Percent Similarity: 79.09% Conservative: 4
Best Local Similarity: 75.45% Mismatches: 17
Query Match: 12.95% Indels: 6
DB: 15 Gaps: 2
US-10-023-523-8 (1-530) x US-10-227-577-467 (1-405)
QY 57 GluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnGln 76
Db 75 AAGAAGCGGGCTGCCAACTGGAAGACATCTAGCACATCTGTGTGGACATAACACG 134
QY 77 GlyGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLys 96

```

Db 135 GGGGGCCCGCGAGGAGTGGGCGACAGGGTGAGCCGGCTGAACCCGAAGATGCAGAGAAG 194
QY 97 SerArgThrTyrValAlaArgAsnGlyGluPro-GluProThrProVal---ValTyrGI 115
Db 195 TCCCGGACCTATGTGGCAGGAATGGGAGCCCTTGACCAACTCCAKTAGTTCATTGG 254
QY 115 YGluYsGluProSerYsGlyAsp-ProAsnThrGluGluIle-ArgGlnSer-AspGI 134
Db 255 AGAGAAGGAACCCCTCCCAAGGGGATTCCAAAACACAGAAAGATTCGGGCAGAGTTGACGA 314
QY 134 uValGlyAspArgAsphis---ArgArgProGlnGluYsLysIleAlaYsGlyLeuGI 153
Db 315 AGTTCGGAGCAAGGAACCACTTCGAGGCCACAGGAGAGAAAGCCAAAGGTTTGGG 374
QY 153 YLysGluIleThrLeuLeuMet 160
Db 375 GAGGATCAGCTTTGCTGATT 396

RESULT 37
US-10-264-049-1238
; Sequence 1238, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birex et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133PI
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1238
; LENGTH: 725
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (226)..(226)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (290)..(290)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (320)..(320)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (324)..(324)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (331)..(331)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (339)..(339)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (384)..(384)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (546)..(546)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature

```

```

; LOCATION: (553)..(553)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (605)..(605)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (616)..(616)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (622)..(622)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (662)..(662)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (668)..(668)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (698)..(698)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (705)..(705)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (712)..(712)
; OTHER INFORMATION: n equals a,t,g, or c
; US-10-264-049-1238

Alignment Scores:
Pred. No.: 5,17e-17 Length: 725
Score: 318.00 Matches: 74
Percent Similarity: 77.05% Conservative: 20
Best Local Similarity: 60.66% Mismatches: 26
Query Match: 11.77% Indels: 3
DB: 16 Gaps: 0

US-10-023-523-8 (1-530) x US-10-264-049-1238 (1-725)
QY 347 LeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnThrHis 366
Db 334 TTRANGAGGCGACAGATCGAGCACAAATACGACAAATGAACAGCANGAGTACAA 393
QY 367 LeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSer 386
Db 394 CTAAACACAGCAGCTTCTCTTTATATGATAAGTTTGAAGAAATTCAGACTACCATGGCA 453
QY 387 LysSerSerGluValPheThrThrPhelysGlnGluMetGluLysMetThrLysLysIle 406
Db 454 AAAAGCAATGAATCGTTTACCACTTCAGACAGAAATGAAAAGATGACAAAGAAAT 513
QY 407 LysLysLeuGluLysGluThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAla 426
Db 514 AAAAACTGGAAAAAGAAACAAATAATTTGGCGNACCAANGGGAACCAT-AAATAAGCA 572
QY 427 LeuLeuGluMetAlaGluGluLysThrValArgAspLysGluLeuGluGlnVal 446
Db 573 CTTTGCAATCGCTGAAGAGGAGAAACAGTCGCGNATAAAGAGTNCAGAGCCCTTCAATA 632
QY 447 LysIleGlnArg-LeuGluLysLeuCysArgAla-LeuGlnThrGluArgAsnAspLeuA 466
Db 633 AAACGTGGAACGGTTAGAGAACTTGGCANGGCTTTNTTCAACCGAAAGGATGGGCTCA 692
QY 466 sn 466
Db 693 AT 694

```

RESULT 38

US-09-876-143-1445
; Sequence 1445, Application US/09876143
; Publication No. US20040081959A1
; GENERAL INFORMATION:
; APPLICANT: Infogen Inc.
; APPLICANT: EILERTSEN, KENNETH J.
; APPLICANT: PFISTER-GENSKOW, MARTHA
; APPLICANT: CHILDS, LYNETTE
; APPLICANT: FORSYTHE, TODD
; APPLICANT: BISHOP, MICHAEL D.
; TITLE OF INVENTION: IDENTIFICATION AND USE OF MOLECULAR MARKERS INDICATING
; TITLE OF INVENTION: CELLULAR REPROGRAMMING
; FILE REFERENCE: 028040-0202
; CURRENT APPLICATION NUMBER: US/09/876,143
; CURRENT FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/209,874
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 1744
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1445
; LENGTH: 981
; TYPE: DNA
; ORGANISM: Bovine
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(981)
; OTHER INFORMATION: n is a, c, g, or t
US-09-876-143-1445

Alignment Scores:

Pred. No.:	1,89e-16	Length:	981
Score:	313.00	Matches:	74
Percent Similarity:	62.88%	Conservative:	9
Best Local Similarity:	56.06%	Mismatches:	46
Query Match:	11.58%	Indels:	4
DB:	12	Gaps:	2

US-10-023-523-8 (1-530) x US-09-876-143-1445 (1-981)

QY	29	GlyProGlySerSerGlnAlaProArgLysProGlyAlaGlnAlaArgThrAlaGln	48
DB	26	GGTGGCGCGCGCTCTAGAACTAGTGGATGCCCC---GGGTGCGAGGANTT-CGGCAGCAG	81
QY	49	SerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGluAspIleLeuSer	68
DB	82	GCTNAGCCCTCCGNGATGINTCTGAGAGTTGANCNCNCANCTGGATGACATCCTCANN	141
QY	69	ThrTyrCysValAspAsnAsnGlnGlyProGlyGluAspGlyAlaGlnGlyGluPro	88
DB	142	ACATACTGCGTGGACAAACAGNANNNGGCGCCAGGTNAGGATGTGGCANNAGGNGAGCCT	201
QY	89	AlaGluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsnGlyGluProGlu	108
DB	202	GCTGACCCGAGATGTCNAGAGNCCNGACCTATNCCTCANGANTGGNAGCCTGAG	261
QY	109	Pro---ThrProValTyrGlyGlyGluProSerLysGlyAspProAsnThrGlu	127
DB	262	CCAGATACTCCAGTAGTCNNNTGNNAGAGGAANNCTCCAAGGGGAGCGCGCCCGGTC	321
QY	128	GluIleArg-GlnSerAspGluValGlyAspArgAspHisArgArgProGlnGlyLys	147
DB	322	TAGATCTGTACCAAGNATGAAGTTTGAGACCCGAGNACCACCCAAAGCCNCGNCGANTATA	381
QY	147	sLysAlaLysGlyLeuGlyLysGluIleThrLeu	158
DB	382	GAAANNCTANGTGTGGGAANGGAATTAACCTTG	415

RESULT 39

US-10-425-114-18551
; Sequence 18551, Application US/10425114
; Publication No. US20040034888A1

GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 18551
; LENGTH: 1035
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3088-030-A1_FLI
US-10-425-114-18551

Alignment Scores:

Pred. No.:	9,89e-16	Length:	1035
Score:	304.50	Matches:	78
Percent Similarity:	55.75%	Conservative:	48
Best Local Similarity:	34.51%	Mismatches:	77
Query Match:	11.27%	Indels:	23
DB:	13	Gaps:	6

US-10-023-523-8 (1-530) x US-10-425-114-18551 (1-1035)

QY	273	GluGlnHisAsnGluArgAsnSerLysLeuArgGlnGlnAsnMetGluLeuAlaGluArg	292
DB	12	GAGCAGAGGGTTGAGTGCAATTCCTCAGCTAGAGAGAGACAATATG---TTGAGAGATAAA	68
QY	293	LeuLysLysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAspLysValPheLys	312
DB	69	CTCAAAGACATTCGTCGATCAGTATAACATTACTACGACGAATAATGCTCACCACCAATTGAAA	128
QY	313	HisLysAspLeuGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMetLeu	332
DB	129	GAGAAATGATGGAACCTTGAGCTTGCTGATCTGAGACITCAACACATCAAGAG-----	182
QY	333	LysGluAlaGluGluArgHisGlnArgGluLysAspPheLeuLysGluAlaValGlu	352
DB	183	AAGGTGCTCAGCAACATACCCAAATGCAA-----TTGTAT	218
QY	353	SerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLysGlnGlnLeuAla	372
DB	219	GCGGAGCAGATTTCTCAGCTTAAG---ACTACTGAGAGAACCTGCGGTGCACTAGCT	275
QY	373	LeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluValPhe	392
DB	276	TCGGATGGAGAAAGATTTTCAGCACITTCAGGATGCCCTGTCAAAAAGCAATGAAGTCTTT	335
QY	393	ThrThrPheLysGlnMetGluLysMetThrLysLysIleLysLysLeuLysGlu	412
DB	336	GAAACTTCAACAGCAGAGATGGAAAAATGATTTGGTGATATAAAATCTTAAAGAGGAC	395
QY	413	ThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeuGluMetAlaGlu	432
DB	396	AATGAATTTCTCAGGGGAAAATGTCAGAACTCAGATATTGCTTGTGAAGCTCATTTGAA	455
QY	433	GluLysThrValArgAspLysGluLeuGlnValLysIleGlnArgLeuGlu	452
DB	456	GAGCGTAGCTAACGAAGAAGCAAAATAGAGAAATTTGAAAAACCAAAAGGACACTTGAG	515
QY	453	LysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgValGlnAspLeu	472
DB	516	TCCTCTGTCTGATCATACTACGCGCAAGAGAAACAA-----	551
QY	473	SerAlaGlyGlnGlySerLeuThrAspSerGlyProGluArgProGluGlyPro	492

Db 552 -----GGCGCTCGCCAGTATTCAGACGCC-----CCTTTAGCCCAAGAGACATGCCA 602
Qy 493 GlyAlaGlnAlaProSer 498
Db 603 CCGACAAGTCAAGATCT 620

RESULT 40
US-09-927-597-3
; Sequence 3, Application US/09927597
; Publication No. US2003032018A1
; GENERAL INFORMATION:
; APPLICANT: Malik, Fady
; APPLICANT: Berard, Christophe
; APPLICANT: Freedman, Richard
; APPLICANT: Craven, Andrew
; APPLICANT: Sakowicz, Roman
; APPLICANT: Hartman, James
; TITLE OF INVENTION: Human smooth muscle myosin heavy chain
; FILE REFERENCE: CYP0018
; CURRENT APPLICATION NUMBER: US/09927597
; CURRENT FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 5937
; TYPE: DNA
; ORGANISM: Human
US-09-927-597-3

Alignment Scores:
Pred. No.: 4,85e-13 Length: 5937
Score: 282.50 Matches: 153
Percent Similarity: 38.48% Conservative: 101
Best Local Similarity: 23.18% Mismatches: 229
Query Match: 10.46% Indels: 178
Gaps: 27

US-10-023-523-8 (1-530) x US-09-927-597-3 (1-5937)

Qy 3 SerProGlyGlnProGluAla-----GlyProGluGlyAlaGlnGluArgPro 18
Db 3969 GCTCCAGGACACCCAGGAGCTGCTTCAAGAGAAGAACCCGCGAGAAGCTCAA---CGTGTC 4025

Qy 19 SerGlnAlaAlaProAlaValGluAlaGluGlyProGlySerGlnAlaProArgLys 38
Db 4026 TACGAAGCTGCCAGCT-----CGAGGAGGCGGACAGCTTCCTCA 4070

Qy 39 ProGluGly---AlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGlu 57
Db 4071 CCAGCTGGAGGAGGAGATGGAGGCGCAAGCAAGCACTGGAGCGCCACATCTCCACTCTCAA 4130

Qy 58 GluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrcysValAspAsnAenGlnGly 77
Db 4131 CATCAGCTCTCCGACTCGAAGAAGAAGAGCTTCAGAGACTTTCGCAGCACCTGGAAGCTCT 4190

Qy 78 GlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluPro----- 91
Db 4191 GGAAGAGGGAAGAAGAGGTTCCAGAGGAGATCGAGAACTCACCCAGCAGTACGAGGA 4250

Qy 92 -----GluAspAlaGluLysSerArgThrTyrcysValAla 102
Db 4251 GAAGCGCGCCGCTTATGATAAAGTGGAAAAGACCAAGAACAGAGCTTCAGCAGGAGTGA 4310

Qy 103 ArgAsn-----GlyGluProGluProThrProValValTyrcysGlyGluLysGlu 118
Db 4311 CGACCTGGTGTGTTGATTGGACACACCGCGCACTCGTGTCCAACTGGAAAAGAGCA 4370

Qy 119 Pro-----SerLysGlyAspProAsnThrGluGluIleArgGlnSer 132
Db 4371 GAGGAATTTGATCAGTTGTTAGCGGAGGAGAAAACATCTCTTCCAAATACGC----- 4424

Qy 133 AspGluValGly-AspArgAspHisArgProGlnGluLysLysAlaLysGlyLe 152
Db 152 -----GATGAGAGGAGCAGAGCTGAGCAGAGCCAGGAGGAGAG----- 4464

Qy 152 uGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLys 172
Db 4465 -----GAAACCAA 4472

Qy 172 sLeuAlaAlaLeuCysLysLysTyrcysAlaGluLeuLeuGluHisArgAsnSerGlnLys 192
Db 4473 GGCCCTGTCCCTGGCTCGGCGCCCTTGAAGAGAGCTTGAAGCCCAAGAGAGAACTCGAGCG 4532

Qy 192 sGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArg 212
Db 4533 GACCACAAATATGCTCAAGCGGAAATCGAAGACCTGTCTAGCTCCAAAGATGAGCTGGG 4592

Qy 212 gGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLe 232
Db 4593 CAAGAACTGCTCATGAGCTGGAGAGTCCAAAGCGGCGCTCGAGACCCAGATGGAGGAGAT 4652

Qy 232 uGlnArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGluGluGlu 252
Db 4653 GAAGACGCGAGCTGGAAGAGCTGGAGAGCAG-----CTGCAAGCCAGGAGAGCGCAA 4706

Qy 252 uLysArgLysGluValThr-----SerHisPheGlnValThrLeuAsnAs 267
Db 4707 ACTGCGGCTGGAAGTCAATGTCAGCGCTCAAGGCGCAGTTCGAA-----AGGGA 4757

Qy 267 pIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMe 287
Db 4758 TCTCAAGCCCGGAGCAGCAGCAATGAGGAGAAAGAGGAGGCAACTGCGAGACAGCTTCA 4817

Qy 287 tGlu-----LeuAlaGluArgLeu 294
Db 4818 CGAGTATGAGCGGAACCTGGAGAGCAGCGAAGCAACGTCCTCGGAGCTGCGAGCAA 4877

Qy 294 sLysLeuIleGlu-----GlnTyrcysGluLeu----- 302
Db 4878 GAAGAAGCTGGAAGGAGGAGCTGAAAGACCTGAGGCTTCAGGCGGCACTGCGCATCAAGGG 4937

Qy 303 -ArgGluGluHisIleAspLysValPheLys-----HisLysAspLeuGlnGlu 318
Db 4938 GAGGAGGAGGAGCCATCAAGCAGCTACGCAAACTGCGAGCTCAGATGAAGAGCTTTCAAAG 4997

Qy 318 nGlnLeuValAspAlaLys----- 324
Db 4998 AGAGCTGGAAGATGCCCTGCTCCAGAGATGAGATCTTTCGACAGCAAGAGAAATGA 5057

Qy 325 -----LeuGlnAlaGlnGluMetLeuLysGluAl 335
Db 5058 GAAGAAAGCCAGAGCTTGAAGAGCAGACCTATGAGCTTCAAGAGAGACCTCGCGCGCGC 5117

Qy 335 aGluGluArgHisGlnArg-----GluLysAspPheLeuLysGluAlaValGlu 352
Db 5118 TGAGAGGCTCGCAAAACAGCGGACCTCGAGAGGAGGAACTGGCAGAGGAGCTGGCCAG 5177

Qy 352 uSerGlnArgMetCysGluLeuMetLysGlnGlnThrHisLysLysGlnGlnLeuAl 372
Db 5178 TAGCCTGTGCGGAAGAAACGCACTCCAGAGCAGAAAGCGCCCTGGAGGCGCGAGCTGC 5237

Qy 372 aLeuTyrcysGluLysPheGluPheGlnAsnThrLeuSerLysSerSerGluValPhe 392
Db 5238 CCAGCTGGAGGAGGAGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5292

Qy 392 eThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLysLysLysGlu 412
Db 5293 -----CGGTCGCGCAAGACCAACAGCAGCGCGCGAGCTCAGCAAGCA 5336

Qy 412 uThrThrMetTyrcysArgSerArgTrpGluSerSerAsnLysAlaLeuLeuGluMet 431
Db 5337 GCTGGCCACAGAGCGCAGCAGCGCCCAAGAAATGAGAGTCCCGGAGAGCTCGAGCG 5396

Qy 431 aGluGluLysThrValArgAspLys-----GluLeuGluGly----- 443
Db 5397 GCAGAAACAGAGAGCTCCCGAGCAAGCTCCAGAGATGAGGAGGCGCGTCAAGTCCAAGTT 5456

Qy 444 -----LeuGlnValIleGlnArgLeuGluLysLeu----- 454
 Db 5457 CAAGTCACCATCCGCGCTGGAGGCCAAGATTGCACAGCTGGAGGAGCAGGTGAGCA 5516
 Qy 455 ----CysArgAlaLeuGlnThrGluArgAsnAspLeu-----AsnLysArgValG1 470
 Db 5517 GGAGCCAGAGAGAAACAGCGCGCCACCAAGTCCTGAAGCAGAAAGCAAGAGCTGAA 5576
 Qy 470 nAspLeuSerAlaGlyGlnGlnGlySerLeuThrAsp-----SerG1 484
 Db 5577 GGAATCTT-GCTCCAGGTGAGAGCAGCGCCAGCAAGATGGCGGACAGTACAAAGAGCAGG 5635
 Qy 484 yProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrG1 504
 Db 5636 CAGAAAGGCAATGCCAGGTCAAGCAGCTCAAGAGCGAGCTGGAGGAGCAGAGAGG 5695
 Qy 504 uAlaProCysTyProGlyAlaProSerThrGluAlaSerGlyGln----- 519
 Db 5696 AGTCCAGCGCATCAAGCGCCAAACCGCAGGAAGCTGCAGCGGAGCTGGATCAGGCCACGG 5755
 Qy 520 -----ThrGlyPro-----GlnGluProThrSerAlaArg 529
 Db 5756 AGACCAAGCGCCATCGCGCGCGAGGTGAAGCAGCTCAAGAGCAGCTCAGCGAGG 5813

RESULT 41

US-10-171-311-161
 ; Sequence 161, Application US/10171311
 ; Publication No. US20030087270A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Schlegel, Robert
 ; APPLICANT: Chen, Yan
 ; APPLICANT: Zhao, Xumei
 ; APPLICANT: Monahan, John
 ; APPLICANT: Kamatkar, Shubhangi
 ; APPLICANT: Glatt, Karen
 ; APPLICANT: Ganavarapu, Manjula
 ; APPLICANT: Hoerish, Sebastian
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
 ; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
 ; TITLE OF INVENTION: OF CERVICAL CANCER
 ; FILE REFERENCE: MRI-035
 ; CURRENT APPLICATION NUMBER: US/10/171,311
 ; CURRENT FILING DATE: 2002-06-12
 ; PRIOR APPLICATION NUMBER: US 60/298,159
 ; PRIOR FILING DATE: 2001-06-13
 ; PRIOR APPLICATION NUMBER: US 60/298,155
 ; PRIOR FILING DATE: 2001-06-13
 ; PRIOR APPLICATION NUMBER: US 60/335,936
 ; PRIOR FILING DATE: 2001-11-14
 ; NUMBER OF SEQ ID NOS: 238
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 161
 ; LENGTH: 6861
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-171-311-161

Alignment Scores:
 Pred. No.: 5,75e-13 Length: 6861
 Score: 282.50 Matches: 153
 Percent Similarity: 38.48% Conservative: 101
 Best Local Similarity: 23.18% Mismatches: 229
 Query Match: 10.46% Indels: 178
 DB: 15 Gaps: 27

US-10-023-523-8 (1-530) x US-10-171-311-161 (1-6861)

Qy 3 SerProGlyGlnProGluAla-----GlyProGluGlyAlaGlnGluArgPro 18
 Db 4036 GCTCAGGACACCCAGGAGCTGCTCAGAGAAACCCGCGCAGAGGTCAA---CGTGC 4092
 Qy 19 SerGlnAlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLys 38

Db 4093 TAGCAGCTGCCCGAGCT-----GGAGGAGGAGCGAAACAGCTTCAAGA 4137
 Qy 39 ProGluGly---AlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGlu 57
 Db 4138 CCAGCTGGACGAGGAGATGGAGCCCAAGCAGAACCTGGAGCGCCACATCTCCACTCTCAA 4197
 Qy 58 GluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrcysValAspAsnAsnGlnGly 77
 Db 4198 CATCCAGCTCTCGAGCTCGAAGAAGAAGCTGCAGAGCTTTGCCAGCACCGTGGAGCTCT 4257
 Qy 78 GlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluPro----- 91
 Db 4258 GGAAGAGGGGAAGAAGAGTTCCAGAAGGAGATCGAAGAACTCACCCAGCAGTACGAGGA 4317
 Qy 92 -----GluAspAlaGluLysSerArgThrTyrcysValAla 102
 Db 4318 GAAGCGCGCGCTTATGATAAAGTGAAGAAACCAAGAACAGGCTTCAGCAGGAGCTGA 4377
 Qy 103 ArgAsn-----GlyGluProGluProThrProValValTyrcysGlyGlu 118
 Db 4378 CGACCTGGTGTGATTGTTGACCAACAGCGGCAACTCGTGTCCAACCTGGAAAGAGCA 4437
 Qy 119 Pro-----SerLysGlyAspProAsnThrGluGluLeuArgGlnSer 132
 Db 4438 GAGGAAATTTGATCAGTTGTTAGCCGAGGAGAAACATCTCTTCCAAATACGC----- 4491
 Qy 133 AspGluValGly-AspArgAspHisArgArgProGlnGluLysLysLysAlaLysGlyLe 152
 Db 4492 -----GGATGAGAGGACAGAGCTGAGCAGAGAGCGCAGGAGAAG----- 4531
 Qy 152 uGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGlu 172
 Db 4532 -----GAAACCAA 4539
 Qy 172 sLeuAlaAlaLeuCysLysLysTy-AlaGluLeuGluHisArgAsnSerGln 192
 Db 4540 GGCCTGTCTCCCTGGCTCGGCGCTTGAAGAGCGCTTGAAGCAAGAGAGAACTCGAGCG 4599
 Qy 192 sGlnMetLysLeuGlnLysGlnSerGlnLeuValGlnGlnLysAspHisLeuArg 212
 Db 4600 GACCAACAAATGCTCAAAAGCCGAAATGAAGACCTGTGTGCTCAGCTCAAGATGACGTGG 4659
 Qy 212 gGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLe 232
 Db 4660 CAAGAAGTCCATGAGCTGGAGAGTCCAAAGCGGCGCTTGAAGAGAGAGAGAGAT 4719
 Qy 232 uGlnArgHisAsnArgSerLeuLysGluGlyValGlnArgAlaArgGluGluGlu 252
 Db 4720 GAAGACGACGCTGGAAGAGCTGAGGACGAG-----CTGCAAGCCACGAGGAGCCAA 4773
 Qy 252 uLysArgLysGluValThr-----SerHisPheGlnValThrLeuAsnAs 267
 Db 4774 ACTGCGCTGGAAGTCAACATCAGCGCTCAGGCGCTCAAGGCGCAGTTTCGAA----- 4824
 Qy 267 pIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMe 287
 Db 4825 TCTCCAAAGCCCGGACGACGACAATGAGAGAGGAGGAGGAGGAGGAGGAGGAGTTCA 4884
 Qy 287 tGlu-----LeuAlaGluArgLeu 294
 Db 4885 CGAGTATGAGACGGAAGTGAAGAGCGAGGAAAGCAACGTGCGCTCGCAGCTGCAGCAA 4944
 Qy 294 sLysLeuIleGlu-----GlnTyrcysLeu----- 302
 Db 4945 GAAGAAGCTGGAAGGGGACCTGAAGACCTGGAGCTTACGGCGGACTCTGCCATCAAGGG 5004
 Qy 303 -ArgGluGluHisIleAspLysValPheLys-----HisLysAspLeuGlnG1 318
 Db 5005 GAGGGAGGAGGAGCATCAAGCAGCTACGCAAACTGCAGGCTCAGATCAAGGACTTTCAAAG 5064
 Qy 318 nGlnLeuValAspAlaLys----- 324

```
Db 5065 AGAGTGAAGATGCCCTGCTCCAGAGATGAGATCTTTCACACGCCAAAGAGATGA 5124
Qy 325 -----LeuGlnAlaGlnGluMetLeuLysGluAl 335
Db 5125 GAAGAAAGCCAGAGCTTGGAGAGACCTCATCAGCTACAGAGAGACCTCGCCGCCG 5184
Qy 335 aGluGluArgHisGlnArg-----GluLysAspPheLeuLysGluAlaValG1 352
Db 5185 TGAGAGGGCTCGCAACAACAGCGGACCTCGAGAGAGAGAACTGGCAGAGCTGGCCAG 5244
Qy 352 uSerGlnArgMetCysGluLeuMetLysGlnGlnThrHisLeuLysGlnGlnLeuAl 372
Db 5245 TAGCTGTCCGAGAGAAACGACCTCCAGAGAGAGAGCGCGCTGAGAGCCCGGATCGC 5304
Qy 372 aLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluValP 392
Db 5305 CCAGCTGGAGAGAGAGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5359
Qy 392 eThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLysLysG1 412
Db 5360 -----CGGGTCCGCAAAAGCCACACAGCAGCGCCGAGCAGCTCAGCAACGA 5403
Qy 412 uThrThrMetTyrArgSerArgTTPGluSerSerAsnLysAlaLeuLeuGluMet---A1 431
Db 5404 GCTGCCACAGAGCGCAGCAGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5463
Qy 431 aGluGluLysThrValArgAspLys-----GluLeuGluGly----- 443
Db 5464 GCAGAAACAGAGAGCTCCGAGCAAGCTCCACAGAGAGAGAGAGAGAGAGAGAGAG 5523
Qy 444 -----LeuGlnValLysLysLysLysLysLysLysLysLysLysLysLysLys 454
Db 5524 CAAGTCCACCATCGCGCGCTGGAGGCGCAAGATTGCAAGCTGGAGAGAGAGAGAG 5583
Qy 455 ---CysArgAlaLeuGlnThrGluArgAsnAspLeu-----AsnLysArgValG1 470
Db 5584 GGAGGCCAGAGAGAAACAGCGCGCCACCAAGCTCGCTGAAGCAGAGAAACAGAGAG 5643
Qy 470 nAspLeuSerAlaGlyGlyGlnGlySerLeuThrAsp-----SerG1 484
Db 5644 GGAATCTT-CTGCAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5702
Qy 484 yProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrG1 504
Db 5703 CAGAGAAAGCAATGCCAGGTCCAGAGCTCAGAGAGAGAGAGAGAGAGAGAGAGAG 5762
Qy 504 uAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGln----- 519
Db 5763 AGTCCAGCGCATCAACGCCAACCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5822
Qy 520 -----ThrGlyPro-----GlnGluProThrSerAlaArg 529
Db 5823 AGAGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5880

RESULT 42
US-10-341-434-102
; Sequence 102, Application US/10341434
; Publication No. US20030215835A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes
; FILE REFERENCE: 9U 204 205 R1
; CURRENT APPLICATION NUMBER: US/10/341,434
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/348,164
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/348,119
; PRIOR FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 102
; LENGTH: 6861
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (89)..(6004)
; OTHER INFORMATION:
US-10-341-434-102

Alignment Scores:
Pred. No.: 5,75e-13 Length: 6861
Score: 282.50 Matches: 153
Percent Similarity: 38.48% Conservative: 101
Best Local Similarity: 23.18% Mismatches: 229
Query Match: 10.46% Indels: 178
DB: 16 Gaps: 27

US-10-023-523-8 (1-530) x US-10-341-434-102 (1-6861)

Qy 3 SerProGlyGlnProGluAla-----GlyProGluGlyAlaGlnGluArgPro 18
Db 4036 GCTTCAGGACACCCAGGAGCTCTTCAAGAGAAACCCGGCAGAGCTCAA---CGTGT 4092
Qy 19 SerGlnAlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLys 38
Db 4093 TAGGAAGCTGCCAGCT-----GGAGGAGGAGCGGAAACAGCTCACAAG 4137
Qy 39 ProGluGly---AlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGlu 57
Db 4138 CCAGCTGGAGAGAGATGGAGGCCAAGCAGCAACCTGGAGCGGCCACATCTCCACTCT 4197
Qy 58 GluLeuSerArgGlnLeuGluAlaAspPheLeuSerThrTyrCysValAspAsnAsnGlnGly 77
Db 4198 CATCAGCTCTCGACTCGAGAGAGAGCTGCGAGGACTTGCACGACCGCTGGAAGCTCT 4257
Qy 78 GlyProGlyGluAspGlyAlaGlnGluProAlaGluPro----- 91
Db 4258 GGAAGAGGGGAAAGAGAGGTTCCAGAGAGAGATCGAGAACTCACCCAGCAGTACGAGA 4317
Qy 92 -----GluAspAlaGluLysSerArgThrTyrValAla 102
Db 4318 GAAGGCGCGCTTATGATAACTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4377
Qy 103 ArgAsn-----GlyGluProGluProThrProValValTyrGlyGluLysGlu 118
Db 4378 CGACCTGGTTGTGATTGTCACACAGCGGCAACTCGTGTCCAACTGGAAAAAGAGCA 4437
Qy 119 Pro-----SerLysGlyAspProAsnThrGluGluLeuArgGlnSer 132
Db 4438 GAGGAAATTTGATCATGTTGTGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4491
Qy 133 AspGluValGly-AspArgAspHisArgArgProGlnGluLysLysLysAlaLysGlyLe 152
Db 4492 -----GGATGAGAGGACAGAGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAG 4531
Qy 152 uGlyLysGluLeuThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluL 172
Db 4532 -----GAAACCAA 4539
Qy 172 sLeuAlaAlaLeuCysLysLysTyrAlaGluLeuLeuGluHisArgAsnSerGlnL 192
Db 4540 GCGCCCTGCTGGCTGGGCGCCCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4599
Qy 192 sGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeu 212
Db 4600 GACCAACAAAATGCTCAAAGCCGAAGACCTGCTGCTCAAGAGTCAAGAGTCAAGTGG 4659
Qy 212 gGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGlu 232
Db 4660 CAAGAACGCTCCATGAGTGGAGAGAGTCCAGCGGCGCCCTGGAGAGAGAGAGAG 4719
Qy 232 uGlnArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGluGluGlu 252
Db 4720 GAAGAGCGAGCTGGAAGAGAGCTGGAGAGAG-----CTGCAAGCCAGGAGAGAG 4773
```


Db 9229 -----GGATGAGAGGACAGAGCTGAGGAGAGAGCCAGGAGAG----- 9190
Qy 152 uGlyLysGluLeuThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGlu 172
Db 9189 -----GAAACCAA 9182
Qy 172 sLeuAlaAlaLeuCysLysLysTyrAlaGluLeuLeuGluGluHisArgAsnSerGln 192
Db 9181 GGCCCTCTCCCTGCTCGGCGCTTGAAGAGGCTTGAAGAGCCAAAGAGGAACTCGAGCG 9122
Qy 192 sGlnMetLysLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeu 212
Db 9121 GACCAACAAATGCTCAAGCCGAATGGAAGACCTGCTCAGTCCAGATGACGTGG 9062
Qy 212 sGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuSerLeuCysArgGlu 232
Db 9061 CAAGAACCTCAGCTGAGTGGAGAGTCCAAAGCGGCGCTGAGACCCAGATGAGGAGAT 9002
Qy 232 uGlnArgHisAsnArgSerLeuLysGluGluValGlnArgAlaArgGluGluGlu 252
Db 9001 GAAGACGAGCTGGAAGAGCTGGAGGAGAG-----CTGACAGCCAGGAGAGCCAA 8948
Qy 252 uLysArgLysGluValThr-----SerHisPheGlnValThrLeuAsnAs 267
Db 8947 ACTGGCTGGAAGTCAACATGTCAGGCGCTCAAGGGCCAGTTCGAA-----AGGGA 8897
Qy 267 PileGlnLeuMetGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMe 287
Db 8896 TCTCCAGCCCGGAGCGAGCAATGAGGAGAGAGAGGAGCACTGCAGAGACGCTTCA 8837
Qy 287 tGlu-----LeuAlaGluArgLeu 294
Db 8836 CGAGTATCAGACGAACTGGAGAGAGCGAAGCAACGCTGCCCTGCGAGCTGCAGCAA 8777
Qy 294 sLysLeuLeu-----GlnTyrGluLeu----- 302
Db 8776 GAAGAACTGGAAGGAGCACTGAAGACCTTGAGCTTCAGCGCGACTCTGCCATCAAGG 8717
Qy 303 -ArgGluGluHisLysLysValPheLys-----HisLysAspLeuGlnG 318
Db 8716 GAGGAGAGACCACTCAAGCAGCTACGCAAACTCCAGGCTCAGATGAGGACTTCAAG 8657
Qy 318 nGlnLeuValAspAlaLys----- 324
Db 8656 AGAGCTGAAGATGCGCTGCTCCAGAGATGAGATCTTGCCACAGCCAAAGAGATGA 8597
Qy 325 -----LeuGlnAlaGlnGluMetLeuLysGluAl 335
Db 8596 GAAGAAACCAAGACTTGAAGAGAGCACTATGCTACAGAGACCTCGCGCGCG 8537
Qy 335 aGluGluArgHisGlnArg-----GluLysAspPheLeuLeuLysGluAlaValG 352
Db 8536 TGAGAGGCTCGCAACAAACGAGGACCTCGAAGAGAGAACTGCGCAGAGGAGCTGGCCAG 8477
Qy 352 uSerGlnArgMetCysGluLeuMetLysGlnGluThrHisLysLysGlnGlnLeuAl 372
Db 8476 TAGCCTGTGCGGAAGAAACGCACTCCAGAGAGAGAGCGCGCTGAGGCGCGGATCGC 8417
Qy 372 aLeuTyrThrGluLysPheGluPheGlnAsnThrLeuSerLysSerSerGluValPh 392
Db 8416 CCAGCTGAGAGGAGCTGAGAGGAGGAGGAGGAGCAACATGGAGCCCATGAGCGAC----- 8362
Qy 392 eThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLysGluGlu 412
Db 8361 -----CGGTGTCGCAAAAGCCACACAGCAGCGCGCGAGCTCAGCAACGA 8318
Qy 412 uThrThrMetTyrArgSerArgTyrGlnSerSerAsnLysAlaLeuLeuMet---Al 431
Db 8317 CTGGCCACAGAGCGCCAGCCGCCCCAGAAATGAGAGTCCCGGAGGAGCTCGAGCG 8258
Qy 431 aGluGluLysThrValArgAspLys-----GluLeuGluGly----- 443
Db 8257 GCAGAAACAGGAGCTCCCGAGCAAGCTCCAGAGATGAGGAGGCGCGCTCAAGTCCAAATT 8198

Qy 444 -----LeuGlnValLysLeuGlnArgLeuGluLysLeu----- 454
Db 8197 CAATCCACCATCGCGCTGGAGGCCAAGATTGCACAGTAGAGGAGGAGTCGAGCA 8138
Qy 455 -----CysArgAlaLeuGlnThrGluArgAsnAspLeu-----AsnLysArgValG 470
Db 8137 GGAGCCAGAGAGAAACAGCGCGCCACCAAGTCGTGAAGCAGAAAGACAAGAGCTCAA 8078
Qy 470 nAspLeuSerAlaGlyGlyGlnGlySerLeuThrAsp-----SerG 484
Db 8077 GCAATCTT-GCTGAGCTGGAGGAGCGCAAGATGGCCGAGCGTGTGATGAGGAGGAG 8019
Qy 484 ProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrG 504
Db 8018 CAGAAAGGCAATCCAGGCTCAGAGCAGCTCAGAGGAGGAGGAGGAGGAGG 7959
Qy 504 uAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGln----- 519
Db 7958 AGTCCAGCGCATCAACCGCAACCGCAGGAAGCTGCAGCGGAGCTGATGAGGAGGAG 7899
Qy 520 -----ThrGlyPro-----GlnGluProThrSerAlaArg 529
Db 7898 AGAGCAACAGAGCCATGGCGCGGAGGAGTGAAGCACTCAGAGCAAGCTCAGGCGAG 7841

RESULT 44

US-09-927-597-1
; Sequence 1, Application US/09927597
; Publication No. US20030032018A1
; GENERAL INFORMATION:
; APPLICANT: Malik, Fady
; APPLICANT: Beraud, Christophe
; APPLICANT: Freedman, Richard
; APPLICANT: Craven, Andrew
; APPLICANT: Sakowicz, Roman
; APPLICANT: Hartman, James
; TITLE OF INVENTION: Human smooth muscle myosin heavy chain
; FILE REFERENCE: CYTO018
; CURRENT APPLICATION NUMBER: US/09/927,597
; CURRENT FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5835
; TYPE: DNA
; ORGANISM: Human
US-09-927-597-1

Alignment Scores:
Pred. No.: 5,728-13 Length: 5835
Score: 281.50 Matches: 147
Percent Similarity: 39.08% Conservative: 101
Best Local Similarity: 23.15% Mismatches: 224
Query Match: 10.42% Indels: 164
DB: 10 Gaps: 25

US-10-023-523-8 (1-530) x US-09-927-597-1 (1-5835)

Qy 3 SerProGlyGlnProGluAla-----GlyProGluGlyAlaGlnGluArgPro 18
Db 3969 GCTCCAGGACACCCAGGAGCTGCTTCAAGAGAAACCCGCGCAGAGCTCAA---CGTGT 4025
Qy 19 SerGlnAlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLys 38
Db 4026 TAGAAGCTTCGCCAGCT-----GGAGGAGGAGCGGACAGCTGCAAGA 4070
Qy 39 ProGluGly---AlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGlu 57
Db 4071 CCAGCTGGAGGAGGAGTGGAGGCGCAAGCAGAACCTGGAGCGCCACACTCTCCACTCTCAA 4130
Qy 58 GluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnGlnGly 77
Db 4131 CATCCAGCTCTCCGACTCGAAGAGAGAGCTGCGAGGACITTTGCCAGCACCGTGGAGCTCT 4190

Alignment Scores:

Pred. No.: 6,99e-13 Length: 6900
 Score: 281.50 Matches: 147
 Percent Similarity: 39.06% Conservative: 101
 Best Local Similarity: 23.15% Mismatches: 224
 Query Match: 10.42% Indels: 164
 Ds: 15 Gaps: 25

US-10-023-523-8 (1-530) x US-10-171-311-163 (1-6900)

QY 3 SerProGlyGlnProGluAla-----GlyProGluGlyAlaGlnGluArgPro 18
 Db 4036 GCTCAGACACCCAGGAGCTCTTCAAGAAAGAACCGCGCAGAGCTCAA---CGTGTC 4092
 QY 19 SerGlnAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLys 38
 Db 4093 TACGAAGCTGCCAGCT-----GGAGGAGGCGGAAACAGCCTGCCAAGA 4137
 QY 39 ProGluGly---AlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGlu 57
 Db 4138 CCAGCTGACAGAGAGATGGAGCCCAACAGCACTGGAGCGCCACATCTCCACTCTCAA 4197
 QY 58 GluLeuSerArgGlnLeuGluAspLeuSerThrTyrCysValAspAsnGlnGly 77
 Db 4198 CATCAGCTCTCCGACTCGGAAGAAAGCTGCAGGACTTTGCCAGCACCGCTGGAGCTCT 4257
 QY 78 GlyProGluGluAspGlyAlaGlnGlyProAlaGluPro----- 91
 Db 4258 GGAAGAGGGGAAGAGAGGTTCCAGAGAGAGATCGAGAACTTCCACCGAGTACGAGGA 4317
 QY 92 -----GluAspAlaGluLysSerArgThrTyrValAla 102
 Db 4318 GAAGCGCGCGCTTATGATAACTGGAAGAGACCAAGAACAGGCTTCAGCAGGAGCTGGA 4377
 QY 103 ArgAsn-----GlyGluProGluProThrProValValTyrGlyGluGlu 118
 Db 4378 CGACCTGGTTGTGATTTGGACACCCAGCGCAACTCGTCCAACTTGGAAAGAGCA 4437
 QY 119 Pro-----SerLysGlyAspProAsnThrGluGluLeuArgGlnSer 132
 Db 4438 GAGGAATTTGATCATGTTAGCGGAGGAGAAACATCTTCCAAATACG----- 4491
 QY 133 AspGluValGly-AspArgAspHisArgProGlnGluLysLysLysAlaLysGlyLe 152
 Db 4492 -----GGATGAGAGGAGCAGAGCTGAGGCAAGCCAGCGGAGAG----- 4531
 QY 152 uGlyLysGluLeuThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGlu 172
 Db 4532 -----GAAACCA 4539
 QY 172 sLeuAlaAlaLeuCysLysLysTyrAlaGluLeuGluGluHisArgAsnSerGlnL 192
 Db 4540 GGCCTGCTCCCTGGCTGGCGCCCTTGAAGAGCGCTTGAAGCCAAAGAGGAACTCGAGCG 4599
 QY 192 sGlnMetLysLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuAr 212
 Db 4600 GACCACAAATGCTCAAGCCGAATGGAAGACCTGTGTAGCTTCCAGGATGACGTGGG 4659
 QY 212 gGlyLysSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLe 232
 Db 4660 CAAGAAGCTCCATGAGCTGGAGAGTCCAAAGCGGCGCTTGGAGCCCAAGATGGAGGAGAT 4719
 QY 232 uGlnArgHisAsnArgSerLeuLysGluGlyValGlnArgAlaArgGluGluGlu 252
 Db 4720 GAAGACGCGCTGGAGAGCTGGAGGAG-----CTGAAGCCAGGAGGAGGCCAA 4773
 QY 252 uLysArgLysGluValThr-----SerHisPheGlnValThrLeuAsnAs 267
 Db 4774 ACTCGGCTGGAGTCAACATCGCGCGCTCAAGGCGCAGTTGAA-----AGGA 4824
 QY 267 pIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMe 287
 Db 4825 TCTCAAGCCCGGAGCAGCAATGAGGAGAGAGGAGGCACTGCGAGACAGACTTCA 4884

QY 287 tGlu-----LeuAlaGluArgLeuLys 294
 Db 4885 CGAGTATGAGACGGAACCTGGAAGAGACGAGGAAAGCAAGTGCCTTGGCAGCTGCAGCAAA 4944
 QY 294 sLysLeuLleGlu-----GlnTyrGluLeu 302
 Db 4945 GAAGAAGCTGGAGGGGACCTGAAAGACCTGAGGCTTCAGGCGGACTGTCCATCAAGGG 5004
 QY 303 -ArgGluGluHisLleAspLysValPheLys-----HisLysAspLeuGlnL 318
 Db 5005 GAGGGAGGAGGAGCTACAGCAGCTACGAAACTGAGGCTCAGATGAAGGACTTTCAAG 5064
 QY 318 nGlnLeuValAspAlaLys----- 324
 Db 5065 AGAGCTGGAGATGCCGTCCTCCAGAGATGAGATCTTCCACAGCAAGAGAGATGA 5124
 QY 325 -----LeuGlnAlaGlnGluMetLeuLysGluAl 335
 Db 5125 GAAGAAAGCAAGAGCTTGAAGACGACCTCATGACGCTACAAGAGGACCTCGCGCGCC 5184
 QY 335 aGluGluArgHisGlnArg-----GluLysAspPheLeuLysGluAlaValG 352
 Db 5185 TGAGAGGCTCGCAACACAGCGGACCTCGAAGAGGAGGAACTGGCAGAGAGCTGGCCAG 5244
 QY 352 uSerGlnArgMetCysGluLeuMetLysGlnGlnThrHisLeuLysGlnLeuAl 372
 Db 5245 TAGCTCTCGGAAAGAACGCACTCCAGGACGAAAGCGCCCTGGAGCGCGGATCGC 5304
 QY 372 aleuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerGluValP 392
 Db 5305 CCAGCTGGAGGAGGAGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5359
 QY 392 eThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLysLysLysLys 412
 Db 5360 -----CGGTCGCGCAAGCCACACAGCAGCGCGGAGGAGCTCAGCAAGCA 5403
 QY 412 uThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeuGluMet---Al 431
 Db 5404 GTTGGCCACAGAGCGGAGCGCGCCCAAGAAATGAGAGTGCCTCGGCGAGCAGCTCGAGCG 5463
 QY 431 aGluGluLysThrValArgAspLys-----GluLeuGluGly----- 443
 Db 5464 GCAGAACAGGAGCTCCGAGGAGAGCTCCAGAGATGGAGGGGCGCGTCACTCCAGT 5523
 QY 444 -----LeuGlnValLysLysLysLysLysLysLysLysLysLysLysLysLys 454
 Db 5524 CAAGTCCACCATCGCGCGCTGGAGGCGCAAGATTGCACAGCTGGAGGAGGAGGAGGAG 5583
 QY 455 -----CysArgAlaLeuGlnThrGluArgAsnAspLeu-----AsnLysArgValG 470
 Db 5584 GGAGGCGCAGAGAGAAACAGCGCGCCACCAAGTCGCTGAAGCAGAGAGAGAGAGAGAG 5643
 QY 470 nAspLeuSerAlaGlyGlyGlnGlySerLeuThrAsp-----SerG 484
 Db 5644 GGAATCTT-GCTGAGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5702
 QY 484 yProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrG 504
 Db 5703 CAGAGAAAGCGCAATGCGAGGCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5762
 QY 504 uAlaProCysTyrProGlyAlaProSerThrGluAlaSerGly 518
 Db 5763 AGTCCCGAGCGCATCACGCGCAACCGCAGGAGAGCTGCGAGCGGA 5805

RESULT 46

US-09-354-456-1602
 ; Sequence 1602, Application US/09954456
 ; Patent No. US20020115057A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Young, Paul
 ; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
 ; TITLE OF INVENTION: Sets

FILE REFERENCE: 689290-76
CURRENT APPLICATION NUMBER: US/09/954,456
CURRENT FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: US/60/233,617
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/60/234,052
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,923
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,134
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,637
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,638
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,711
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,720
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,840
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,863
PRIOR FILING DATE: 2000-09-27
NUMBER OF SEQ ID NOS: 2276
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1602
LENGTH: 3388
TYPE: DNA
ORGANISM: Homo sapiens
US-09-954-456-1602

Alignment Scores:
Pred. No.: 5,26e-13 Length: 3388
Score: 278.50 Matches: 146
Percent Similarity: 38.90% Conservative: 101
Best Local Similarity: 22.99% Mismatches: 225
Query Match: 10.31% Indels: 164
DB: 9 Caps: 25

US-10-023-523-8 (1-530) x US-09-954-456-1602 (1-3388)

QY 3 SerProGlycInPro-----GluAlaGlyProGluGlyAlaGlnGluArgPro 18
Db 1290 GCTCCAGGACCCAGGAGTTGTTCAAGAAAGAACCCCGCAGAGCTCAA---CGTGT 1346
QY 19 SerGlnAlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLys 38
Db 1347 TACGAGCTGGCCAGCT-----GGAGGAGCGGGAACAGCTGCAAGA 1391
QY 39 ProGluGly---AlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgValSerGlu 57
Db 1392 CCAGCTGGACGAGGATGAGGCGCAAGCAGAACCTGGAGCGCCACATCTCCACTCTCAA 1451
QY 58 GluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrcysValAspAsnGlnGly 77
Db 1452 CATCCAGCTCTCGACTCGAAGAGAGGCTGCGAGACTTTGCCAGCACGTTGGAAGCTCT 1511
QY 78 GlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluPro----- 91
Db 1512 GGAAGAGGGAAGAAGAGGTTCAGAGGAGATCGAGAACCTCACCCAGCAGTACGAGGA 1571
QY 92 -----GluAspAlaGluLysSerArgThrTyrrValAla 102
Db 1572 GAAGGCGCGCTTATGATATAAATGAAAGAGCAAGACAGGCTTCAGCAGGAGCTGA 1631
QY 103 ArgAsn-----GlyGluProGluProThrProValValTyrrGlyGluLysGlu 118
Db 1632 CGACCTGGTTGTTGATTGTCACACAGCGGCAACTCGTGTCCAACTGGAAAGAGCA 1691
QY 119 Pro-----SerLysGlyAspProAsnThrGluGluLeuArgGlnSer 132
Db 1692 GAGGAATAATTGATCAGTTGTTAGCCGAGGAGAAACATCTCTTCCAAATACGC----- 1745

QY 133 AspGluValGly-AspArgAspHisArgArgProGlnGluLysLysLysAlaLysGlyLe 152
Db 1746 -----GGATGAGAGGACAGAGCTGAGCAGCAAGCAGGAGAAAG----- 1785
QY 152 uGlyLysGluIleThrLeuLeuValGlnThrLeuAsnThrLeuSerThrProGluGlu 172
Db 1786 -----GAAACCA 1793
QY 172 sleuAlaAlaLeuCysLysLysTyrrAlaGluLeuGluHisArgAsnSerGlnly 192
Db 1794 GGCCTGTCTCTGGCTCGGGCCCTTGAAGAGCCCTTGAAGCCAAAGAGGAACTCGAGCG 1853
QY 192 sGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuAr 212
Db 1854 GACCAACAAATGCTCAAGCGCAATGGAAGACCTGCTCAGCTCCAAAGGATGAGCTGG 1913
QY 212 gGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLe 232
Db 1914 CAAGAAGCTCCATGAGCTGGAGAGTCCCAAGCGGCCCTGGAGACCCAGATGGAGAGAT 1973
QY 232 uGlnArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGluGluGlu 252
Db 1974 GAAGACGCGAGCTGGAGAGCTGGAGAGCAG-----CTGCAAGCTCGGAGGCGCAA 2027
QY 252 uLysArgLysGluValThr-----SerHisPheGlnValThrLeuAsnAs 267
Db 2028 ACTGGCGCTGGAAGTCAACATGTCAGCGCTCAAGGCCAGTTCGAA-----AGGA 2078
QY 267 pIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMe 287
Db 2079 TCTCAAGCCCGGACGAGCAGATGAGAGAGAGGAGGCACTGCGAGAGAGCTTCA 2138
QY 287 tGlu-----LeuAlaGluArgLeu 294
Db 2139 CGAGTATGAGCGGAACCTGGAAGACGAGCGAAACGACGCTGCTGGCAGCTGCGACAA 2198
QY 294 sLysLeuLeGlu-----GlnTyrrGluLeu----- 302
Db 2199 GAAGAAGCTGGAAGGGGACCTGAAAGACCTGAGGCTTCAGGCGGACTGTCATCAAGG 2258
QY 303 -ArgGluGluHisIleAspLysValPheLys-----HisLysAspLeuGlnGlu 318
Db 2259 GAGGAGAGAGCCTCAAGAGCTACGCAAACTCGAGCTCAGATGAAGGACTTTCAAG 2318
QY 318 nGlnLeuValAspAlaLys----- 324
Db 2319 AGAGCTGGAGATGCCGTGCTCCAGAGATGAGATCTTCCACAGCAGCAAGAGATGA 2378
QY 325 -----LeuGlnAlaGlnGluMetLeuLysGluAl 335
Db 2379 GAAGAAGCCCAAGAGCTTGAAGCAGACCTCATGCGCTACAAAGAGGAGCTCGCGCGC 2438
QY 335 aGluGluArgHisGlnArg-----GluLysAspPheLeuLeuLysGluAlaVal 352
Db 2439 TGAGAGGCTCGCAAAACAGCGGACCTCGAAGAGGAACTGGCAGAGGAGCTGGCCAG 2498
QY 352 uSerGlnArgMetCysGluLeuMetLysGlnGlnThrHisLysLysGlnGlnLeuAl 372
Db 2499 TAGCCTGTGGAGAGAAACGCACTCCAGAGCAGAAAGCGCCCTGGAGGCGCGATCGC 2558
QY 372 aleuTyrrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluValPh 392
Db 2559 CCAGCTGAGGAGGAGCTGGAGGAGGAGCAGGCAACATGGAGGCCATGAGCGAC----- 2613
QY 392 eThrThrPheLysGlnGluMetGluLysMetThrLysLysIleLysLysLeuGluLysGl 412
Db 2614 -----CGGTCCGCAAAAGCCACAGCAGCGCCGAGAGCTCAGCAACGA 2657
QY 412 uThrThrMetTyrrArgSerArgTrpGluSerSerAsnLysAlaLeuLeuGluMet---Al 431
Db 2658 GCTGCCACAGAGCGCAGCAGCGCCCAAGAGATGAGAGTCCCGGCGAGCAGCTCGAGCG 2747
QY 431 aGluGluLysThrValArgAspLys-----GluLeuGluGly----- 443

Db 2718 GCAGAACAGGAGCTCGGAGCAAGTCCACGAGATGGAGGGGCGCTCAAGTCCAGTT 2777
QY 444 -----LeuGlnValIleGlnArgLeuGluLeu----- 454
Db 2778 CAGATCCACCATCGGCGCTGGAGCCAGATTGCACAGCTGGAGGAGCAGCTGCAGCA 2837
QY 455 -----CysArgAlaLeuGlnThrGluArgAsnAspLeu-----AsnIysArgValGI 470
Db 2838 CGAGGCCAGAGAGAAACAGGCAGCCACCAAGTCGCTGAAGCAGACAGAAAGACAGAAAGCTGAA 2897
QY 470 nAspLeuSerAlaGlyGlyGlnGlySerLeuThrAsp-----SerGI 484
Db 2898 GGAATCTT-GCTGCAGTGGAGGACGAGCGCAAGATGGCCGAGCAGTACAGGACGAG 2956
QY 484 yProGluArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGI 504
Db 2957 CAGAGAAAGCAATGCCAGGCTCAAGCAGCTCAAGAGGCGAGCTGGAGGAGGCAGAGGAG 3016
QY 504 uAlaProCysTy-ProGlyAlaProSerThrGluAlaSerGly 518
Db 3017 AGTCCCGAGCATCAACGCCAACCCGAGGAAGCTGCAGCGGGA 3059

RESULT 47

US-09-967-768A-245
; Sequence 245, Application US/09967768A
; Patent No. US20020150877A1
; GENERAL INFORMATION:
; APPLICANT: Augustus, Meena
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-72
; CURRENT APPLICATION NUMBER: US/09/967,768A
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/60/236,109
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,034
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,111
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 325
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 245
; LENGTH: 3388
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-967-768A-245

Alignment Scores:
Pred. No.: 5,26e-13 Length: 3388
Score: 278.50 Matches: 146
Percent Similarity: 38.90% Conservative: 101
Best Local Similarity: 22.99% Mismatches: 225
Query Match: 10.31% Indels: 164
DB: 9 Gaps: 25

US-10-023-523-8 (1-530) x US-09-967-768A-245 (1-3388)

QY 3 SerProGlyGlnPro-----GluAlaGlyProGluGlyAlaGlnGluArgPro 18
Db 1290 GCTCCAGGACACCCAGGAGTGTCTCAAGAGAAACCCGCGCAGAGCTCAA---CGTGTC 1346
QY 19 SerGlnAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLys 38
Db 1347 TACGAAGCTGCGCCAGCT-----GGAGGAGGCGGAAACAGCTGCAGAA 1391
QY 39 ProGluGly---AlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGlu 57
Db 1392 CCAGCTGACGAGAGATGGGGCCAGCAGACCTGGAGGGCCACATCTCCACTCTCAA 1451
QY 58 GluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGly 77
Db 1452 CATCCAGCTCTCCGACTCGAAGAGAGCTGCAGGACTTTGCCAGCACCGCTGGAAGCTCT 1511

QY 78 GlyProGlyGluAspGlyAlaGlnGlyProAlaGluPro----- 91
Db 1512 GGAAGAGGGGAAGAGAGGTTCCAGAGAGGATCGAGAACCTCACCCAGCAGTACGAGGA 1571
QY 92 -----GluAspAlaGluLysSerArgThrTyrValAla 102
Db 1572 GAAGGGCGCCCTTATGATAAAGTGGAAAGACCAAGAACAGGCTTCAGCAGGAGCTGGA 1631
QY 103 ArgAsn-----GlyGluProGluProThrProValValTyrGlyGluLysGlu 118
Db 1632 CGACCTGTTGTTGATTGGACAACAGCGCAACTCGTCTCCAACTGGAAGAAAGCA 1691
QY 119 Pro-----SerLysGlyAspProAsnThrGluGluLeuArgGlnSer 132
Db 1692 GAGGAATTTGATCAGTTGTTAGCCGAGGAGAAACATCTCTCCAATACGC----- 1745
QY 133 AspGluValGly-AspArgAspHisArgA-gProGlnGlnLysLysLysAlaLysGlyLe 152
Db 1746 -----GGATGAGAGGGACAGAGCTGAGGCGCAAGACCGAGGAGAAG----- 1785
QY 152 uGlyLysGluLeuThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLy 172
Db 1786 -----GAAACCAA 1793
QY 172 sLeuAlaAlaLeuCysLysLysTyrAlaGluLeuLeuGluHisArgAsnSerGlnLy 192
Db 1794 GGCCCTGTCCTCGCTCGGCGCTTGAAGAGGCTTGGAGCCCAAGAGAACTCGAGCG 1853
QY 192 sGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGlnLysAspHisLeuAr 212
Db 1854 GACCAACAAATGCTCAAGCCGAAATGGAAGACCTGCTCAGCTCCAGAGTACGTGGG 1913
QY 212 sGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLe 232
Db 1914 CMAGAACGTCATGAGCTGGAGAGTCCACGCGGCGCTCGAGACCCAGATGGAGGAGAT 1973
QY 232 uGlnArgHisAsnArgSerLeuLysGluGluValGlnArgAlaArgGlnGluGluG 252
Db 1974 GAAGACGCGAGCTGGAAGAGCTGGAGAGCGAG-----CTGCAAGCCTCGGAGGAGCGCAA 2027
QY 252 uLysArgLysGluValThr-----SerHisPheGlnValThrLeuAsnAs 267
Db 2028 ACTGCGCTGGAAGTCAACATGCGAGCGCTCAAGGCGCAGTTCGAA-----AGGGA 2078
QY 267 pileGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMe 287
Db 2079 TCTCCAGCCCGGAGCAGCAGCAAGTACGAGAGAGGAGGCAACTGCAGACACAGCTTCA 2138
QY 287 tGlu-----LeuAlaGluArgLeuLy 294
Db 2139 CGAGTATGAGCGGAACCTGGAAGACGAGCGAAACGAGCTGCTTGGCAGCTGCAGCAA 2198
QY 294 sLysLeuIleGlu-----GlnTyrGluLeu----- 302
Db 2199 GAAGAAGCTGGAGGGGACCTGAAAGACCTTCAGGCGGACTTCGCCATCAAGGG 2258
QY 303 ArgGluGluHisIleAspLysValPheLys-----HisLysAspLeuGlnG 318
Db 2259 GAGGAGGAGGAGCCATCAAGCAGCTACGCAAACTCGAGGCTCAGATGAAGGAGCTTCAA 2318
QY 318 nGlnLeuValAspAlaLys----- 324
Db 2319 AGAGCTGGAAGATGCCCTGCTCCAGAGATGAGATCTTCCACAGCAAGAGAGATGA 2378
QY 325 -----LeuGlnGlnAlaGlnGlnMetLeuLysGluAl 335
Db 2379 GAAGAAAGCCCAAGAGCTTGGAAAGCAGACCTCATGCACTACAGAGGAGACCTCGCGCGCG 2438
QY 335 aGluGluArgHisGlnArg-----GluLysAspPheLeuLeuLysGluAlaValGI 352
Db 2439 TGAGAGGCTCGCAACAAAGCGGACCTCGAGAGGAGGAACTGGCAGAGGAGCTGGCCAG 2498

QY	352	uSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLeuLysGlnGlnLeuAl	372
Db	2499	TAGCCCTGTGCGGGAAGAAAGCGCACTCCAGGACGAGCAAGCCCGCTTGGAGCCCGGATGCG	2558
QY	372	aLeuTYrThrGluLysPheGluPheGlnAsnThrLeuSerLysSerSerGluValPhe	392
Db	2559	CCAGCTGGAGGAGGAGCTGGAGGAGGAGCGGCGCAACATGGAGGCCATGAGGCGAC----	2613
QY	392	eThrThrPheLysGlnGlnMetGluLysMetThrLysLysLysLysLysLysLysGln	412
Db	2614	-----CGGTCGCGAAAGCCACACAGCAGGCGGAGCGAGCAGCTCAGCAACGA	2657
QY	412	uThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeuGluMet----	431
Db	2658	GCTGGCCACAGAGCGCAGCAAGCGCCAGAGAATGAGAGTGCCCGGCACACAGCTCGAGCG	2717
QY	431	agluGluLysThrValArgAspLys-----gluLeuGluGly-----	443
Db	2718	GCAGAAACAAGGAGCTCCGAGGCAAGTCCACAGAGATGGAGGGGGCGCGTCAAGTCCAAGTT	2777
QY	444	-----LeuGlnValLysIleGlnArgLeuGluLysLeu-----	454
Db	2778	CAAGTCCACCATCGCGCGCTGGAGGCCAAGATTGCACAGCTGGAGGACAGGTGAGCA	2837
QY	455	----CysArgAlaLeuGlnThrGluArgAsnAspLeu-----	470
Db	2838	GGAGGCCAGAGAGAAACAGCAGCCACCAAGTCGCTGGAAGCAGAAAGACAAGACTGAA	2897
QY	470	nAspLeuSerAlaGlyGlyGlnGlySerLeuThrAsp-----	484
Db	2898	GGAAATCTT-GCTGCGAGTGGAGGACGAGCGCAAGATGCCCGAGCAGTACAAGGAGCAGG	2956
QY	484	yProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrCl	504
Db	2957	CAGAGAAAGCAATGCAGCGGTCAACAGCTCAAGGAGCAGCTGGAGGAGGAGGAGGAGG	3016
QY	504	uAlaProCysTyrProGlyAlaProSerThrGluAlaSerGly	518
Db	3017	AGTCCAGCGGCATCAACGCGCAACCCGCGAGGAAGTGCAGCGGGA	3059

RESULT 48

US-09-954-531-988

; Sequence 988, Application US/09954531

; Patent No. US20020165180A1

; GENERAL INFORMATION:

; APPLICANT: Weaver, Zee

; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Can

; TITLE OF INVENTION: Gene Sets

; FILE REFERENCE: 689290-77

; CURRENT APPLICATION NUMBER: US/09/954,531

; CURRENT FILING DATE: 2002-05-02

; PRIOR FILING DATE: US/60/233,133

; PRIOR FILING DATE: 2000-09-18

; PRIOR APPLICATION NUMBER: US/60/234,009

; PRIOR FILING DATE: 2000-09-20

; PRIOR APPLICATION NUMBER: US/60/234,034

; PRIOR FILING DATE: 2000-09-20

; PRIOR APPLICATION NUMBER: US/60/234,509

; PRIOR FILING DATE: 2000-09-22

; PRIOR APPLICATION NUMBER: US/60/234,567

; PRIOR FILING DATE: 2000-09-22

; NUMBER OF SEQ ID NOS: 1392

; SOFTWARE: Patent in version 3.0

; SEQ ID NO 988

; LENGTH: 3389

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-954-531-988

Alignment Scores:		
Pred. No.:	5,26e-13	Length:
Score:	278.50	Matches:
Percent Similarity:	38.90%	Conservative:
		101

Best Local Similarity:	22.99%	Mismatches:	225
Query Match:	10.31%	Indels:	164
DB:	9	Gaps:	25

US-10-023-523-8 (1-530) x US-09-954-531-988 (1-3388)

Qy	3	SerProGlyGlnPro-----	-----GluAlaGlyProGluGlyAlaGlnGluArgPro	18
		:::		
Db	1290	GCTCAGACACCGAGAGTTGCTT	CAAGAAGAACCCGGCAGAGCTCAA	---CGTGC 1346
Qy	19	SerGlnAlaProLaValGluAlaGluGlyProGlySerSerGlnAlaProArgLys	38	
		:::		
Db	1347	TACGAAGCTGGCAGCT-----	---GGAGAGAGCGGAACACGCTGCAAGA	1391
Qy	39	ProGluGly--AlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGlu	57	
		:::		
Db	1392	CCAGCTGGACGAGGATGGAGCCACAGACAACTGGAGGCCACATCTCCACTCTCAA	1451	
Qy	58	GluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrcysValAspAsnAsnGlnGly	77	
		:::		
Db	1452	CATCAGCTCTCGAGCTCGAAGAAGCTGCAGAGCTTTCCAGACACCGTGGAGAGCTCT	1511	
Qy	78	GlyProGluAspGlyAlaGlnGlyGluProAlaGluPro-----	---91	
		:::		
Db	1512	GGAAGAGGGAGAGGAGGTTCCAGAGAGAGATCCAGAACCTCACCAGCAGTACGAGGA	1571	
Qy	92	-----GluAspAlaGluLysSerArgThrTyrrValAla	102	
		:::		
Db	1572	GAAGCGCGCGCTTATGATAAACTGAAAGAACCCAGAAACAGGCTTCAGCAGAGCTGGA	1631	
Qy	103	ArgAsn-----GlyGluProGluProThrProValValTyrrGlyGluLysGlu	118	
		:::		
Db	1632	CGACCTGTTGTTGATTGGACAAACCGCGCACTCGTGTCACACCTGGAAAGAAGCA	1691	
Qy	119	Pro-----SerTyrsGlyAspProAsnThrGluGluLeuArgGlnSer	132	
		:::		
Db	1692	GAGGAATTTGATCAGTTGTTACCGGAGAGAAAACATCTCTCCAAATACGC-----	1745	
Qy	133	AspGluValGly-AspArgAspHisArgArgProGlnGluLysLysLysAlaLysGlyLe	152	
		:::		
Db	1746	---GGATCAGAGGGACACAGACTGAGGCACAGACCGAGGAGAAG-----	1785	
Qy	152	uGlyLysGluLeuThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLys	172	
		:::		
Db	1786	-----GAAACCAA	1793	
Qy	172	sLeuAlaAlaLeuCysLysLysTyrrAlaGluLeuLeuGluGluHisArgAsnSerGlnLys	192	
		:::		
Db	1794	GGCCCTGCTCGCTCGGCGCCTTGAAGAGCGCTTGAAGCCAAAGAGAACTCGAGCG	1853	
Qy	192	sGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuAr	212	
		:::		
Db	1854	GACCACAAATGCTCAAGCCCAATGAGACCTGTCTAGCTCCAGAGATGACGTGGG	1913	
Qy	212	gGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLe	232	
		:::		
Db	1914	CAAGAACTCATCATGCTGGAGAGTCCAAAGCGGCGCTTGGAGACCCAGATGAGGAGAT	1973	
Qy	232	uGlnArgHisAsnArgSerLeuLysGluGlyValGlnArgAlaArgGluGluGluGlu	252	
		:::		
Db	1974	GAAGACGAGCTGGAAGAGCTGGAGACAG-----CTGCAAGCTTCGGAGNACGCCAA	2027	
Qy	252	uLysArgLysGluValThr-----SerHisPheGlnValThrLeuAsnAs	267	
		:::		
Db	2028	ACTCGGCTGGAAGTCAACATGACGCGCTCAAGGGCCAGTTCCAA-----AGGGA	2078	
Qy	267	pIleGlnLeuMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluGluGlu	287	
		:::		
Db	2079	TCTCCAGCCGGGACGACGACGATGAGGAGNAGAGGAGGAGCAACTGCGAGACAGCTTCA	2138	
Qy	287	tGlu-----LeuAlaGluArgLeuLys	294	
		:::		
Db	2139	CGAGTATGAGACCGGAACCTGGAAGACGAGCAACGAGACGTCTCGTGCACCTGCACAAA	2198	

Alignment Scores:	
Pred. No.:	5.26e-13
Scores:	278.50
Percent Similarity:	38.50%
Length:	3388
Matches:	146
Conservative:	101

[illegible]

```

QY 294 slysluileGlu-----GlnTyrGluLeu----- 302
Db 2199 GAAGAAGCTGAAGAGGAGCCTGAAGACCTGGAGCTTCAGGCGGACTGTCATCAAGGG 2258
QY 303 -ArgGluGluHisIleAspLysValPheLys-----HisLysAspLeuGlnG1 318
Db 2259 GAGGAGGAGGAGCCTCAAGCAGCTAGCGAACTGCGAGCTCAGATGAGGAGCTTTCAGG 2318
QY 318 nGlnLeuValAspAlaLys----- 324
Db 2319 AGAGCTGGAAGATGCCCGTCCAGATGAGATCTTGGCCACAGCAGCAAGAGATGA 2378
QY 325 -----LeuGlnGlnAlaGlnGluMetLeuLysGluAl 335
Db 2379 GAAGAAGCCAGAGCTTGAAGAGCAGCTCATGCTACAGAGGACCTCGCGCGCG 2438
QY 335 aGluGluAraGHisGlnArg-----GluLysAspPheLeuLeuLysGluAlaValG1 352
Db 2439 TGAGAGGCTCGCAACAGCGGAGCTCGAGAGGAGGAACTGGCAGAGGAGCTGGCGAG 2498
QY 352 uSerGlnArgMetCysGluLeuMetLysGlnGlnGlnGlnHisLysLysGlnGlnLeuAl 372
Db 2499 TAGCCTGTGCGGAAGGAGCACTCCAGGACGAGAAAGCGCGCTGGAGGCGCGGATCGC 2558
QY 372 aLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluValPh 392
Db 2559 CCAGCTGAGGAGGAGCTCGAGGAGGAGCGGCAACATGAGGCGCCATGAGCGAGC 2613
QY 392 eThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLysLysGluLysG1 412
Db 2614 -----CGSGTCCGCAAGCCACACAGCAGCGCGCGAGCTCAGCAAGCA 2657
QY 412 uThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeuGluMet---Al 431
Db 2658 GTGGCCACAGAGCGGAGCAGCGCCAGCAAGATGAGAGTCCCGGAGCAGCTCGAGCG 2717
QY 431 aGluGluLysThrValArgAspLys-----GluLeuGluGly 443
Db 2718 GCAGAAACAGGAGCTCCGAGCAAGCTCCAGAGATGAGGAGGCGCGCTCAAGTCCAAGTT 2777
QY 444 -----LeuGlnValLysIleGlnArgLeuGluLysLeu----- 454
Db 2778 CAAGTCCACATCGCGGCGCTGAGGCGCAAGATTCACAGCTGAGGAGGAGGAGTTCGAGCA 2837
QY 455 ---CysArgAlaLeuGlnThrGluArgAsnAspLeu-----AsnLysArgValG1 470
Db 2838 GGAGGCCACAGAGAAACAGCGCAGCCACCAAGTCGCTGAAGCAGAAAGAGAGCTGAA 2897
QY 470 nAspLeuSerAlaGlyGlyGlyGlySerLeuThrAsp-----SerG1 484
Db 2898 GGAATCTT-TGCTGAGGTGAGGAGCAGCGCAAGATGCGCGACAGTACAAGAGAGCGG 2956
QY 484 yProGluArgArgProGluGlyProGlyValaGlnAlaProSerSerProArgValThrG1 504
Db 2957 CAGAGAAAGGCAATCCAGGCTCAAGCAGCTCAAGAGGAGCAGCTGGAGGAGGAGGAGG 3016
QY 504 uAlaProCysTyrProGlyAlaProSerThrGluAlaSerGly 518
Db 3017 AGTCCAGCGCATCAACGCCAACCCGAGAGAGCTGCGAGCGGGA 3059

RESULT 49
US-09-954-531-1382
; Sequence 1382, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
; APPLICANT: Weaver, Zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; FILE OF INVENTION: Gene Sets
; FILE REFERENCE: 689290-77
; CURRENT APPLICATION NUMBER: US/09/954,531
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/60/233,133

```

```

; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 1392
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1382
; LENGTH: 3388
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-531-1382

Alignment Scores:
Pred. No.: 5,26e-13 Length: 3388
Score: 278.50 Matches: 146
Percent Similarity: 38.90% Conservative: 101
Best Local Similarity: 22.90% Mismatches: 225
Query Match: 10.31% Indels: 164
DB: 9 Gaps: 25

US-10-023-523-8 (1-530) x US-09-954-531-1382 (1-3388)
QY 3 SerProGlyGlnPro-----GluAlaGlyProGluGlyAlaGlnGluArgPro 18
Db 1290 GCTCCAGGACACCCAGGAGTGTCTTCAAGAGAAACCCGCGCAGAGCTCAA---CGTGTC 1346
QY 19 SerGlnAlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLys 38
Db 1347 TACGAGCTGCGCGAGCT-----GGAGGAGGAGCGGAGGAGCTGCAAGA 1391
QY 39 ProGluGly---AlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGlu 57
Db 1392 CCAGCTGAGCAGGAGATGAGGCGCAAGCAGAAACCTGGAGCGGCACATCTCCACTCTCAA 1451
QY 58 GluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnGlnGly 77
Db 1452 CATCCAGCTCTCCGCTCGAGAGAGAGCTGCGAGCTTTCAGGACCGTGGAGAGCTCT 1511
QY 78 GlyProGlyLysAspGlyAlaGlnGlyGluProAlaGluPro----- 91
Db 1512 GGAAGAGGGGAGAGAGAGGAGTTCAGAGAGAGATCGAGAAACCTCACCCAGCAGTACGAGA 1571
QY 92 -----GluAspAlaGluLysSerArgThrTyrValAla 102
Db 1572 GAAGCGCGCGCTTATGATAACTGGAAGAGACCAAGACAGAGCTTCACGAGAGCTGA 1631
QY 103 ArgAsn-----GlyGluProGluProThrProValValTyrGlyGluLysGlu 118
Db 1632 CGACCTGGTTGTTGATTGACCAACAGCGGCAACTCGTGTCCCACTCGAAAAGAGCA 1691
QY 119 Pro-----SerLysGlyAspProAsnThrGluGluLeuArgGlnSer 132
Db 1692 GAGAAATTTGATCAGTTGTTAGCCGAGAGAGAAACATCTCTTCCAAATAGCC----- 1745
QY 133 AspGluValGly-AspArgAspHisArgArgProGlnGluLysLysLysAlaLysGlyLe 152
Db 1746 -----GGATGAGAGGACAGAGCTGAGGCGAGAGCGGAGGAGGAGGAGGAGGAGG 1785
QY 152 uGlyLysGluLeuThrLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluGly 172
Db 1786 -----GAAACCAA 1793
QY 172 sLeuAlaAlaLeuCysLysLysTyrAlaGluLeuLeuGluHisArgAsnSerGlnLys 192
Db 1794 GGCCTGCTCCCTGCTGCGGCCCTTGAAGAGGCGCTTGAAGAGGCGCTTGAAGAGGCGCT 1853
QY 192 sGlnMetLysLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuAr 212

```

Db 1854 GACCAACAAATGCTCAAGCCGAAATGGAAGACCTGGTCAAGTCCAAAGGATGACGTGG 1913
Qy 212 SGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLysCysArgGluE 232
Db 1914 CAAGAACGCTCCATCAGCTGGAGAGTTCGAAGCGGGCCCTGGAGACCCAGATGAGGAGAT 1973
Qy 232 uGlnArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGluGluGlu 252
Db 1974 GAAGACGACCTGGAAGAGCTGGAGGACGAG-----CTCAAGCCTCGGAGGACGCCAA 2027
Qy 252 uArgLysGluValThr-----SerHisPheGlnValThrLeuAsnAs 267
Db 2028 ACTGCGCTGGAAGTCAACATGCAAGCGCTCAAGGGCCAGTTCGAA-----AGGA 2078
Qy 267 pLeGlnLeuGlnMetGluGlnHisGlnArgAsnSerLysLeuArgGlnGlnAsnMe 287
Db 2079 TCCTCAAGCCCGGACGACGAGATGAGGAGAGAGAGGAGGCAACTGCAGAGAGCTTCA 2138
Qy 287 tGlu-----LeuAlaGluArgLeuLys 294
Db 2139 CGAGTATGACGAGAACTGGAAGACGAGCGAAACGAGCTGGCCCTGGCAGCTGCACAAA 2198
Qy 294 sLysLeuLeuGlu-----GlnTyrGluLeu----- 302
Db 2199 GAAGAGCTGGAAGGGGACCTGGAAGACCTGGAGCTTCAGCGGACCTTGCCATCAAGGG 2258
Qy 303 -ArgGluGluHisLysAspLysValPheLys-----HisLysAspLeuGlnG 318
Db 2259 GAGGAGAGGACCATCAAGCAGCTACCAACTGCAGGCTCAGATGAAGCTTCAAG 2318
Qy 318 nGlnLeuValAspAlaLys----- 324
Db 2319 AGAGCTGGAAGATGCCGCTCCAGAGATGAGATCTTTGCCACAGCCAAAGAGAATGA 2378
Qy 325 -----LeuGlnGlnAlaGlnGluMetLeuLysGluAl 335
Db 2379 GAAGAAAGCAAGAGCTTGAAGCAGACCTTCAGCTCAAGAGACCTCCGCCCGC 2438
Qy 335 aGluGluArgHisGlnArg-----GluLysAspPheLeuLeuLysGluAlaValG 352
Db 2439 TGAGAGGCTCGCAACAAAGCGGACCTCGAAGAGGAGAACTGGCAGAGAGTGCCAG 2498
Qy 352 uSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLeuLysGlnGlnLeuAl 372
Db 2499 TAGCTGCTGGGAGAGCAACGATCCAGGACGAGCGCGCTGGAGCCCGGATCGC 2558
Qy 372 aLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluValPh 392
Db 2559 CCAGCTGGAGGAGAGCTGGAGGAGGAGCAGGCAACATGGAGGCCATGAGCCAC----- 2613
Qy 392 eThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLysLysLysG 412
Db 2614 -----CGGGTCCGCAAGCCACACAGCAGGCGCGGAGCTCAGCAACGA 2657
Qy 412 uThrThrMetTyrArgSerArgTrpGluSerSerAsnLysAlaLeuLeuMet---Al 431
Db 2658 GCTGGCCACAGAGCGGACGAGCGCCCAAGAGATGAGATGCGCGGAGCGCTCGAGCG 2717
Qy 431 aGluGluLysThrValArgAspLys-----GluLeuGluGly----- 443
Db 2718 GCAGAACAGGAGCTCGGAGCAAGCTCCACGAGATGGAGGGCGCGCTCAAGTCCAAGTT 2777
Qy 444 -----LeuGlnValLysLysLysLysLysLysLysLysLysLysLysLysLys 454
Db 2778 CAAGTCCACCATCGCGCGCTGGAGGCAAGATGTCAGCTGGAGGAGGAGGAGGAGGAG 2837
Qy 455 ---CysArgAlaLeuGlnThrGluArgAsnAspLeu-----AsnLysArgValG 470
Db 2838 GGAGGCGAGAGAGAAACAGGACGCCCAAGTCCGTGAACGACAGCAAGAGAGCTGAA 2897
Qy 470 nAspLeuSerAlaGlyGlyGlnGlySerLeuThrAsp-----SerG 484
Db 2898 GGAATATCTT-GCTCAGGTGGAGGAGCGAGCGCAAGATGGCCGAGCAGTACAGGAGCAGG 2956

Qy 484 yProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrG 504
Db 2957 CAGAGAGGCAATCCAGGGTCAAGCAGCTCAAGAGCGAGCTGGAGGAGCGAGGAGG 3016
Qy 504 uAlaProCysTyrProGlyAlaProSerThrGluAlaSerGly 518
Db 3017 AGTCCAGCGCATCAACCAACCGCAGGAAAGCTGCAGCGGA 3059

RESULT 50

US-09-873-367C-85
; Sequence 85, Application US/09873367C
; Publication NO. US20030165839A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; APPLICANT: Soppet, Daniel
; APPLICANT: Endress, Gregory
; APPLICANT: Augustus, Meena
; APPLICANT: Ebner, Reinhard
; APPLICANT: Carter, Kenneth
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE REFERENCE: 689290-64
; CURRENT APPLICATION NUMBER: US/09/873,367C
; PRIOR FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,891
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,842
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/244,867
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: U.S. 60/245,084
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 1067
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 85
; LENGTH: 3388
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-367C-85

Alignment Scores:
Pred. No.: 5,268-13 Length: 3388
Score: 278.50 Matches: 146
Percent Similarity: 38.90% Conservative: 101
Best Local Similarity: 22.99% Mismatches: 225
Query Match: 10.31% Indels: 164
DB: 10 Gaps: 25

US-10-023-523-8 (1-530) x US-09-873-367C-85 (1-3388)

Qy 3 SerProGlyGlnPro-----GluAlaGlyProGluGlyAlaGlnGluArgPro 18
Db 1290 GCTCCAGGACACCCAGGAGTTCCTCAAGAGAAACCGCGGAGAGCTCAA---CGTGTC 1346
Qy 19 SerGlnAlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLys 38
Db 1347 TAGAAGCTGGCCAGCT-----GGAGGAGGAGCGGAACAGCTCAAGA 1391
Qy 39 ProGluGly---AlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGlu 57
Db 1392 CCAGCTGGAGGAGGAGATGGAGGCCAAGCAGACACCTGGAGCGCCACATCTCCACTCTCAA 1451
Qy 58 GluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnGlnGly 77
Db 1452 CATCCAGCTCTCCGACTCGAAGAGAGAGCTTCAGGAGCTTTCCAGCACCCGCGGAGGAGCTCT 1511
Qy 78 GlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluPro----- 91
Db 1512 GGAAGAGGAGAGAGAGAGTTCAGAGAGAGATCGAGAACCTCACCCAGCAGTACGAGGA 1571
Qy 92 -----GluAspAlaGluLysSerArgThrTyrValAla 102

Db 1572 GAAGCGCGCGCTTATGATAAATCGAAAGAACCAAGACAGGCTTCAGCAGGAGCTGGA 1631
Qy 103 ArgAsn-----GlyGluProGluProThrProValValTyGlyGluLysGlu 118
Db 1632 CGACCTGGTGTGTTGATTTGGACAACACCGGCAACTCGTGTCCAACTCGAAAGAACGA 1691
Qy 119 Pro-----SerLysGlyAspProAsnThrGluGluLeuLeuArgGlnSer 132
Db 1692 GAGGAATTTGATCAGTTGTAGCGGAGGAGAAACATCTCTTCCAAATACGC----- 1745
Qy 133 AspGluValGly-AspArgAspHisArgArgProGlnGlnLysLysLysLysGlyLe 152
Db 1746 -----GGATGAGAGGGGACAGAGCTGAGCGCAAGACCCAGGAGGAAG----- 1785
Qy 152 uGlyLysGluLeuThrLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLy 172
Db 1786 -----GAAACCAA 1793
Qy 172 sLeuAlaAlaLeuCysLysLysTyAlaGluLeuLeuGluHisArgAsnSerGlnLy 192
Db 1794 GGCCCTGTCCCTGCTCGGCGCTTGAAGAGCGCTTGAAGAGCCCTTGAAGAGCAAACTCGAGCG 1853
Qy 192 sGlnMetLysLeuGlnLysLysGlnSerGlnLeuValGlnLysAspHisLeuAr 212
Db 1854 GACCAACAAATGCTCAAGCGGAAATGGAAGACCTGTGCTCAAGGATGACGTGGG 1913
Qy 212 gGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLysCysArgGluLe 232
Db 1914 CAAGAACCTCCATGAGCTGGAGAAATCCAGCGCGCTTGAAGACCCAGATGGAGGAGAT 1973
Qy 232 uGlnArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGluGluGlu 252
Db 1974 GAAGACGCGAGCTGGAAGAGCTGGAGGACGAG-----CTGCAGCCCTCGAGAGCGCAA 2027
Qy 252 uLysArgLysGluValThr-----SerHisPheGlnValThrLeuAsnAs 267
Db 2028 ACTGCGGCTGGAAGTCAACATGCGAGCGCTCAAGGGCCAGTTCGAA-----AGGA 2078
Qy 267 pIleGlnLeuGlnMetGlnGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMe 287
Db 2079 TCTCAAGCCCGGACGAGCAGATGAGGAGAGGAGGAGCACTGCGAGACAGACAGCTTCA 2138
Qy 287 tGlu-----LeuAlaGluArgLeuLy 294
Db 2139 CGAGTATGAGACGGAACTGGAGAGAGGAGGAGCAACGAGCTGCGCTGCGAGCTGCGCAAA 2198
Qy 294 sLysLeuileGlu-----GlnTyGluLeu----- 302
Db 2199 GAAGAAGCTGGAAGGGGACCTGAAAGACCTGAGCTTCAGGCGGACTCTGCCATCAAGGG 2258
Qy 303 -ArgGluGluHisLysLysValPheLys-----HisLysAspLeuGlnGlu 318
Db 2259 GAGGAGGAGGAGCCATCAAGCAGCTACGCAAACTCGAGGCTCAGATGAGGAGCTTCAAG 2318
Qy 318 nGlnLeuValAspAlaLys----- 324
Db 2319 AGAGCTGGAAGATGCCGTGCTCCAGAGATGAGATCTTTGCCACAGCAAGAGAGATGA 2378
Qy 325 -----LeuGlnGlnAlaGlnGluMetLeuLysGluAl 335
Db 2379 GAAGAAACCCAGAGCTTGGAGAGCAGACCTCATCGAGCTACAGAGGACCTCGCGCGCGC 2438
Qy 335 aGluGluArgHisGlnArg-----GluLysAspPheLeuLeuLysGluAlaValG 352
Db 2439 TGAGAGGCTCGCAACACAGCGGACCTCGAGAGGAGGAACTGGCAGAGGAGCTGGCCAG 2498
Qy 352 uSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLysLysGlnGlnLeuAl 372
Db 2499 TAGCCTGTGCGGAGGAGAAACGCACTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2558
Qy 372 aLeuTyThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluValPh 392
Db 2559 CCAGCTGAGGAGGAGCTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2613

Search completed: June 8, 2004, 13:13:40
Job time : 1129.75 secs

Qy 392 eThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLysLysLysGlu 412
Db 2614 -----CGGGTCCGCAAAAGCCACACAGCAGCGCCGAGCAGCTCAGCAACGA 2657
Qy 412 uThrThrMetTyArgSerArgTrpGluSerSerAsnLysLysLysLysLysLysLysMet 431
Db 2658 GCTGCGCCACAGAGCCGACGCGCCAGAGAAATGAGAGTCCCGGCGAGCAGCTCGAGCG 2717
Qy 431 aGluGluLysThrValArgAspLys-----GluLeuGluGly----- 443
Db 2718 GCAGAACAGGAGCTCCCGAGCAAGCTCCACGAGATGAGGGGGCCGTCAGTCCAAAT 2777
Qy 444 -----LeuGlnValLysLysLysLysLysLysLysLysLysLysLysLysLys 454
Db 2778 CAAAGTCCACCATCGCGCGCTGGAGGCCAAGATTGCACAGCTGGAGGAGCAGGTGAGCA 2837
Qy 455 -----CysArgAlaLeuGlnThrGluArgAsnAspLeu-----AsnLysArgValG 470
Db 2838 GGAGGCCAGAGAGAAACAGGAGCCACCAAGTCCGCTGAAGCAGAAAGCAAGAGCTGAA 2897
Qy 470 nAspLeuSerAlaGlyGlyGlnGlySerLeuThrAsp-----SerG 484
Db 2898 GGAATCTT-GCTGCAGGTGGAGGAGCAGCGCAAGATGGCCGAGCAGTACAAGGAGCAGG 2956
Qy 484 yProGluArgArgProGluGlyProGlyValGlnAlaProSerSerProArgValThrG 504
Db 2957 CAGAGAAAGGCCAATCCAGGCTCAAGCAGCTCAAGAGCAGCTGGAGGAGGAGGAGGAG 3016
Qy 504 uAlaProCysTyProGlyAlaProSerThrGluAlaSerGly 518
Db 3017 AGTCCAGCGCATCAACGCCAACCCAGCAAGAGCTGCAGCGGA 3059

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 8, 2004, 06:37:58 ; Search time 3868.11 Seconds
(without alignments)
4091.647 Million cell updates/sec

Title: US-10-023-523-8
Perfect scores: 2702
Sequence: 1 KSSPGQPEAGPEGAQRPESQ.....APSTASGQTGPQPTSARA 530
Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Zgapop 6.0, Zgapext 7.0
Delop 6.0, Delext 7.0
Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Command line parameters:
-Q=/cgn2_1/USPTO.spool/US10023523/runat_08062004_063749_29950/app_query.fasta_1.1422
-DB=EST -QFMT=FASTAP -SUFFIX=rst -MINWATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=50
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: em_estba:**
2: em_esthum:**
3: em_estin:**
4: em_estm:**
5: em_estov:**
6: em_estpl:**
7: em_estro:**
8: em_htc:**
9: gb_est1:**
10: gb_est2:**
11: gb_htc:**
12: gb_est3:**
13: gb_est4:**
14: gb_est5:**
15: em_estfun:**
16: em_estom:**
17: em_gss_hum:**
18: em_gss_inv:**
19: em_gss_pln:**
20: em_gss_vrt:**
21: em_gss_fun:**
22: em_gss_mam:**
23: em_gss_mus:**
24: em_gss_pro:**
25: em_gss_rtd:**
26: em_gss_png:**
27: em_gss_vrl:**
28: gb_gss1:**

29: gb_gss2:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1291	47.8	2797	11	AK031783 Mus muscu
2	1291	47.8	2832	11	AK044130 Mus muscu
3	1291	47.8	3810	11	AK030100 Mus muscu
4	1279	47.3	798	14	CA328534 UI-M-FYO-
5	1269	47.0	775	14	CF728426 UI-M-HBO-
6	1264	46.8	770	14	CF735205 UI-M-HBO-
7	1250.5	46.3	932	12	BG827157 602751143
8	1243.5	46.0	806	14	CA320886 UI-M-FWO-
9	1235.5	45.7	903	13	BX752507 BX752507
10	1234.5	45.7	779	14	CB247992 UI-M-FIO-
11	1225	45.3	945	9	AL961454 AL961454
12	1209	44.7	725	13	BUT01958 UI-M-FIO-
13	1207	44.7	735	13	BQ769910 UI-M-FIO-
14	1205	44.6	759	14	CF995807 AGENCOURT
15	1203	44.5	817	12	BG761234 602718583
16	1194	44.2	1608	11	AK084639 Mus muscu
17	1193	44.2	2480	11	AK085904 Mus muscu
18	1161.5	43.0	750	14	CA327097 UI-M-FYO-
19	1160	42.9	942	13	BQ714219 AGENCOURT
20	1154	42.7	727	13	BQ770181 UI-M-FIO-
21	1153.5	42.7	716	14	CF728006 UI-M-HBO-
22	1140.5	42.2	819	10	BF783468 602111238
23	1136.5	42.1	4452	11	AK031023 Mus muscu
24	1128	41.7	778	14	CA324134 UI-M-FYO-
25	1127.5	41.7	727	14	CB244426 UI-M-FYO-
26	1116	41.3	699	13	BUT01604 UI-M-FIO-
27	1099	40.7	780	13	BX855901 BX855901
28	1081	40.0	702	9	AL637771 AL637771
29	1078	39.9	799	13	BUS49604 AGENCOURT
30	1070	39.6	698	12	BJ618489 BJ618489
31	1069.5	39.6	1318	12	BM455349 AGENCOURT
32	1061	39.3	861	12	B1102887 60288485
33	1060	39.2	647	14	CD578468 UI-M-FYO-
34	1059	39.2	653	13	BUE13886 UI-M-EWO-
35	1044.5	38.7	722	14	CA510893 UI-R-FJO-
36	1040	38.5	665	9	AL887857 AL887857
37	1040	38.5	735	12	BM963869 UI-M-EEO-
38	1037	38.4	782	13	BUT31392 603947670
39	1037	38.4	862	14	CA793605 AGENCOURT
40	1034	38.3	654	13	BQ389733 NISC_mq09
41	1025	37.9	930	13	BQ950691 AGENCOURT
42	1014.5	37.5	954	13	BUI51934 AGENCOURT
43	991.5	36.7	965	10	BF796069 602258906
44	988	36.6	1611	29	AY405165 Homo sapi
45	985	36.5	620	10	AW412004 uc055f09.y
46	978.5	36.2	970	13	BX370444 BX370444
47	977	36.2	720	13	BUS59309 603478867
48	974	36.0	813	13	BQ421247 AGENCOURT
49	972	36.0	784	14	CB247697 UI-M-FIO-
50	967.5	35.8	796	12	BG975353 602842995

ALIGNMENTS

RESULT 1
AK031783
LOCUS
DEFINITION Mus musculus 11 days embryo head cDNA, RIKEN full-length enriched library, clone:6230424F16 product:hypothetical protein, full insert sequence.
ACCESSION AK031783
VERSION AK031783.1 GI:26327606

AK031783 2797 bp mRNA linear HTC 18-SEP-2003
Mus musculus 11 days embryo head cDNA, RIKEN full-length enriched library, clone:6230424F16 product:hypothetical protein, full insert sequence.

KEYWORDS

HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 Carninci, P. and Hayashizaki, Y.

AUTHORS

High-efficiency full-length cDNA cloning

JOURNAL

Meth. Enzymol. 303, 19-44 (1999)

MEDLINE

99279253

PUBMED

10349636

REFERENCE

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

AUTHORS

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE

Normalization and subtraction of cap-trapper-selected cDNAs to

JOURNAL

prepare full-length cDNA libraries for rapid discovery of new genes

MEDLINE

Genome Res. 10 (10), 1617-1630 (2000)

PUBMED

20499374

REFERENCE

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,

AUTHORS

Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,

TITLE

Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

JOURNAL

Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,

MEDLINE

Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wataniki, M.,

PUBMED

Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,

REFERENCE

Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

TITLE

RIKEN integrated sequence analysis (RISA) system--384-format

JOURNAL

sequencing pipeline with 384 multiplexed sequencer

MEDLINE

Genome Res. 10 (11), 1757-1771 (2000)

PUBMED

20530913

REFERENCE

4 The RIKEN Genome Exploration Research Group Phase II Team and the

AUTHORS

FANTOM Consortium.

TITLE

Functional annotation of a full-length mouse cDNA collection

JOURNAL

Nature 409, 695-690 (2001)

PUBMED

11076861

REFERENCE

5 The FANTOM Consortium and the RIKEN Genome Exploration Research

AUTHORS

Group Phase I & II Team.

TITLE

Analysis of the mouse transcriptome based on functional annotation

JOURNAL

of 60,770 full-length cDNAs

PUBMED

Nature 420, 563-573 (2002)

REFERENCE

6 (bases 1 to 2797)

AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,

TITLE

Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,

JOURNAL

Hayashida, K., Hayatsu, N., Hiramoto, K., Hirao, T., Hirozane, T.,

PUBMED

Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,

REFERENCE

Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,

AUTHORS

Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

source

1. 2797
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="PANTOM DB:6230424F16"
 /db_xref="MGI:2393099"
 /clone="6230424F16"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="11 days embryo"
 29. 1603
 /note="unnamed protein product; hypothetical protein
 (evidence: rSCDS, ProCrest, decoder, NCBI CDS
 Predictor, Longest-ORF)
 putative"
 /codon_start=1
 /protein_id="BAC27547.1"
 /db_xref="GI:26327607"
 /translation="MATRLLEVTGRGGTETASGGRRRSPQKFEIGTWEAR
 IGLGVKADVMCNQANDILQDPSGCGTKKSHLEGDSDFITKRNLVSVFCT
 QKREIIPGRARTPPDQDQSECRNKTGLKEVLLQALNTLSTPEKLAALC
 KKVADLEPSRVQKMTLOKQAFIVKEVHLSKALARSKLESLRELQRL
 NKLLEENMQARERERKEATAQITLNEIQALQEHQDTHNAKLENIJELK
 KXLSQVALRESHIDKFKHLEQQLVDKLTQOTLIKEADEKHQERBFLKEAT
 ESRHYEQMKQEVQLKQSLYMKDFEPQTMAKSNELTTFQENKMTKKKL
 EKETIWRKTNKNAKLLQMAEEKVDRKYKAFQIKLERLEKLCRALQTERNE
 KYEVLKQVSIKADGDLVSPATPCAVLDSFKETSRRTLGHLEARAKSVCEKSAQ
 KPSSGSPAQIGESVD"

CDS

Alignment Scores:
 Pred. No.: 1,398-74 Length: 2797
 Score: 1291.00 Matches: 260
 Percent Similarity: 71.77% Conservative: 68
 Best Local Similarity: 56.89% Mismatches: 67
 Query Match: 47.78% Indels: 62
 DB: 11 Gaps: 4

ORIGIN

US-10-023-523-8 (1-530) x AK031783 (1-2797)
 Qy 19 SerGlnAlaProAlaValGluAlaGluGlyPro-----GlySerSerGlnAlaPro 36
 Db 200 TCTCAAGCAATGATATTCTTCAACATCAAGACCCAGTTGTGTGGCACAACAAAGAA 259
 Qy 37 ArgLysProGluGlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgValSer 56
 Db 260 CATTCACTGAAGGGGATGAAGC-----AGT 286
 Qy 57 GluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysVal---AspAsnAsn 75
 Db 287 GACTTTATTACAAAGAACAGAAATTTGGTGAGCTCAGTATTCTGTACACAGGAGAAAGA 346
 Qy 76 GlnGlyGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGlu 95
 Db 347 GAAGAAATTCCTGGACGAGAAAGCTCGAACAGGTCCTCTGATGGCCAGCAAGATTCAGAG 406
 Qy 96 LysSerArgThrTyrValAlaArgAsnGlyGluProGluProThrProValValTyrGly 115
 Db 407 TGCAGCAGG----- 415
 Qy 116 GluLysGluProSerLysGlyAspProAsnThrGluGluIleArgGlnSerAspGluVal 135
 Db 415 ----- 415
 Qy 136 GlyAspArgAspHisArgArgProGlnGluLysLysAlaLysGlyLeuGlyLysGlu 155
 Db 416 -----AACAAAGAGAGACCTTAGAAGAA 442
 Qy 156 IleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGlnLysLeuAla 175
 Db 443 GTTTTATTACTGATGCAAGCGGTAAACACCTTTTCAACCCCGAGAGAGAGCTGCAGCT 502

FEATURES

Location/Qualifiers

/db_xref="MGI:2409425"
 /db_xref="taxon:10090"
 /clone="A830093C13"
 /issue_type="cortex"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="10 days neonate"
 64_1638
 /note="unnamed protein product; hypothetical protein
 (evidence: rscds, prodict, decoder, NCBI CDS
 Predictor, longest-ORF)
 putative"
 /codon_start=1
 /protein_id="BAC31791.1"
 /db_xref="GI:26336212"
 /translation="MATLEVTGRGGTTEASEGGRRSSPPQKEIGTMEPAR
 IGLGVKADVMVNSQANDILOHQPDSGGTTHKSLGSDGSDPITKNLSVSVCT
 QKREIPIGRNRTGPPGQDSECSRNKXTLKEVLLMLQALNTLSPKLAALC
 KYADLLEESNVQMKILQKQKQIVKEKHQSEHSAIKAILARSKLSLCELRQK
 NKLKEENMQQAEEREEKEATAFOITLNEQALQEOHDHNAKLQENELGELK
 KXLEFYALREEDHDKVFKELQQLVDKLTQTLQLEKADREHOREFELKEAT
 ESRHYEQMKQEQVLKQSLYMDKFEFTWAKSNELFTFROEMKMTKKLKL
 EKETIIRTKWNNKALLQMAEKTVDKEYKAFQIKLERLEKLCALQETERNELNE
 KVELKEOVSLKADGDVLPATQPCAVLDSFKETSRRTLGMHLEARKSVCEKSAQ
 KPSSSGSPAQIESVD"

CDS

Alignment Scores:
 Pred. No.: 1,428-74 Length: 2832
 Score: 1291.00 Matches: 260
 Percent Similarity: 71.77% Conservative: 68
 Best Local Similarity: 56.89% Mismatches: 67
 Query Match: 47.78% Indels: 62
 DB: 11 Gaps: 4

US-10-023-523-8 (1-530) x AK044130 (1-2832)

ORIGIN

19 SerGlnAlaProAlaValGluAlaGluGlyPro-----GlySerSerGlnAlaPro 36
 235 TCTCAAGCAAAATGATATCTTCAACATCAAGACCCAGATTGTGTGGCAACAATAAGAAA 294
 37 ArgLysProGluGlyAlaGlnAlaGlnSerGlyAlaLeuArgAspValSer 56
 295 CATTCACTGAGAGGGATGAGGC-----AGT 321
 57 GluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysVal---AspAsnAsn 75
 322 GACTTATTACAAAGAACAGAAATTTGGTGAGCTCAGTATTCTGTACACAGGAGAAAAGA 381
 76 GlnGlyGlyProGluGlyAlaGlnGlyProAlaGluProGluAspAlaGlu 95
 382 GAAGAAATTCCTGGACGAGAGCTCGAACAGGTCTCTGATGCCGACGAGATTCAGAG 441
 96 LysSerArgThrTyrValAlaArgAsnGlyGluProGluProThrProValValTyrGly 115
 442 TGCAGCAGG----- 450
 116 GluLysGluProSerLysGlyAspProAsnThrGluGluIleArgGlnSerAspGluVal 135
 450 ----- 450
 136 GlyAspArgAspHisArgArgProGlnGluLysLysAlaLysGlyLeuGlyLysGlu 155
 451 -----AACAAAGAGAGAGACCTTTAGGAAAGAA 477
 156 IleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAla 175
 478 GTTTTATTACTGATCAGCGCTTAACACCTTTCAACCCAGAGGAGAGCTGGCAGCT 537
 176 LeuCysLysLysTyrAlaGluLeuGluGluHisArgAsnSerGlnLysGlnMetLys 195
 538 CTCCTGAGAAATATGCTGATCTCTGAGAAAGAGCAGGAATTTTCAGAAACAAATGAG 597
 196 LeuLeuGlnLysLysGlnSerGlnLeuValGlnGlnLysAspHisLeuArgGlyGluHis 215

598 ATTCTGCAGAGAGAGCAAGCCAGATTGTGAAGAGAGAAAGTTTCCTTCAGAGTGAACAC 657
 216 SerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHis 235
 658 AGCAAGGCCATCTTGGCAAGAACCAATGGAATCTCTTTGAGGGAACTTCAGCGTCA 717
 236 AsnArgSerLeuLysGluGlyValGlnArgAlaArgGluGluGluLysArgLys 255
 718 AATAAGACCTTTAAGAGGAGGAGAAATATGCAGCAGGACGAGAGAGAGAGACGAGTAAA 777
 256 GluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHis 275
 778 GAAGCAACAGCATTTCCAGATAACTCTAAATGAAATCCAAAGCTCAGTTGGAACAACAT 837
 276 AsnGluArgAsnSerLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLys 295
 838 GACATCCACAATCCCAACTGCGACAGGAGAAACATTTGAACTGGGAGAGAAAGTTGAAGA 897
 296 LeuIleGluGlnTyrGluLeuArgGluGluHisIleAspLysValPheLysHisLysAsp 315
 898 CTTATTGAGCAGTATGCACCTAAGGGAAGACGATATTGATAAGTATTCAACACACAAGAA 957
 316 LeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAla 335
 958 TTGCACACACGCTTGTGTGATGCCAACTTCAGCAACACACACAGCTGATAAAGAGCT 1017
 336 GluGluArgHisGlnArgGluLysAspPheLeuLysGluAlaValGluSerGlnArg 355
 1018 GATGAAAAACATCAGAGAGAGAGAGAGTTTATTATAAAGAGCAACAAGATCCAGGAC 1077
 356 MetCysGluLeuMetLysGlnGlnGlnThrHisLeuLysGlnGlnLeuAlaLeuTyr 375
 1078 AATATGACAAATGAAACAGCAGAGAGTACACATTAACACAGCAGCTTCTCTTTATG 1137
 376 GluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluValPheThrPhe 395
 1138 GATAAATTTTCAAGATTTCCAGACTACTATGGCAAAAGCAATGAACATTTTACACCTTC 1197
 396 LysGlnGluMetGluLysMetThrLysLysIleLysLysLeuLysGluLysGluThrMet 415
 1198 AGCAGGAATGGAAGATGACAAAGAAATTAATAAATCTGGAAGAAACAAATATA 1257
 416 TyrArgSerArgTrpGluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThr 435
 1258 TGGCTACCAATATGGAAACCAATATAAAGCATTCTGCAAGTGGCCGAGAGAAAAC 1317
 436 ValArgAspLysGluLeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLysCys 455
 1318 GTCCGTGATAAGAGTACAGGCTTTTCAATATAAAGCTGGAACGGTTAGAGAACTGTGC 1377
 456 ArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgValGlnAspLeu 472
 1378 AGGCTCTTCAGACAGAGAGAAATGAGCTCAACAGAGAGGTGGAAGCTCTG 1428

RESULT 3
 AK030100 3810 bp mRNA linear HTC 18-SEP-2003
 LOCUS AK030100
 DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:493244K18 product:hypospadias protein, full insert sequence.
 ACCESSION AK030100
 VERSION AK030100.1 GI:263226082
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 REFERENCE 1 Carninci, P. and Hayashizaki, Y.
 AUTHORS High-efficiency full-length cDNA cloning
 TITLE Meth. Enzymol. 303, 19-44 (1999)
 JOURNAL 99279253
 MEDLINE


```

10349636
2  CARNINCI, P., SHIBATA, Y., HAYATSU, N., SUGAHARA, Y., SHIBATA, K.,
REFERENCE ITOH, M., KONNO, H., OKAZAKI, Y., MURAMATSU, M., AND HAYASHIZAKI, Y.
AUTHORS Normalization and subtraction of cap-trapper-selected cDNAs to
TITLE Prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
3  SHIBATA, K., ITOH, M., AIZAWA, K., NAGASAKI, S., SASAKI, N., CARNINCI, P.,
REFERENCE KONNO, H., AKIYAMA, J., NISHI, K., KITSUNAI, T., TASHIRO, H., ITOH, M.,
AUTHORS SUMI, N., ISHII, Y., NAKAMURA, S., HAZAMA, M., NISHINE, T., HARADA, A.,
YAMAMOTO, R., MATSUMOTO, H., SAKAGUCHI, S., ISEGAMI, I., KASHIWAGI, K.,
FUJIWAKE, S., INOUE, K., TODAWA, Y., IZAWA, M., OHARA, E., WATANIKI, M.,
YONEDA, Y., ISHIKAWA, T., OKAWA, K., TANAKA, T., MATSUURA, S., KAWAI, J.,
OKAZAKI, Y., MURAMATSU, M., INOUE, Y., KIRA, A., AND HAYASHIZAKI, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multipillar sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
4  THE RIKEN GENOME EXPLORATION RESEARCH GROUP PHASE II TEAM AND THE
REFERENCE FANTOM CONSORTIUM.
AUTHORS Functional annotation of a full-length mouse cDNA collection
TITLE Nature 409, 685-690 (2001)
JOURNAL 5
REFERENCE The FANTOM Consortium and the RIKEN Genome Exploration Research
AUTHORS Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 3810)
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, N.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M., and Hayashizaki, Y.
DIRECT SUBMISSION
TITLE Submitted (18-JUL-2001) Yoshihide Hayashizaki, The Institute of
JOURNAL Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Sushiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.
FEATURES Location/Qualifiers
source
1..3810
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:4932441K18"
/db_xref="MGI:2392716"
/db_xref="taxon:10090"
/clone="4932441K18"
/sex="male"
/tissue_type="testis"

```

```

/clone lib="RIKEN full-length enriched mouse cDNA library"
/dev stage="adult"
62..1636
/note="unnamed protein product; hypothetical protein
(evidence: rSCDS, ProCrest, decoder, NCBI CDS
Predictor, Longest-ORF)
putative"
/codon_start=1
/protein_id="BAC26785.1"
/db_xref="GI:26326083"
/translation="NATLEEVTRGGGTGTEASEGGRRRRPPPOKFEIGTMEEAR
ICGLGVKAMVNCNSOANDILOHODPSCGTTKXISLEGDEGSDRTITRNRLVSVFCT
OEKREIPGEARTGPPDQGDSECSRNGTCKLGEVLLAOLANTLSTPEKLAALC
KKTADLEENRNQVQEMIKLQKQAIYKVKHQSESKALIALSLKESLCRELRQH
KLIIEYALREEDHDKVFKELQOOLVDKQOTQLIKEDSKHQRERFLKEAT
ERHRYQKQKQVQLKQQLYMDPFEFTWAKNSLEFLTFRQEMERKTKLKEK
EKETITWTKWNNKALIOAEKTVDEKVKAFQIKLERLEKLCALQTERNELNE
KVELKQVSIKADQDLVSPATQPCAVLDSEFKETSRRTLGMLHLEASAKSVCEKSAQ
KPSSSGSPAQIESVD"
ORIGIN
Alignment Scores:
Pred. NO.: 2,03e-74 Length: 3810
Score: 1291.00 Matches: 260
Percent Similarity: 71.77% Conservative: 68
Best Local Similarity: 56.89% Mismatches: 67
Query Match: 47.78% Indels: 62
DB: 11 Gaps: 4
US-10-023-523-8 (1-530) x AK030100 (1-3810)
QY 19 SerGlnAlaAlaProAlaValGluAlaGluGlyPro-----GlySerSerGlnAlaPro 36
Db 233 TCTCAAGCAAAATGATATTCTTCAACATCAAGACCCAGTTGGTGGCCAGCTAAGAAA 292
QY 37 ArgLysProGluGlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSer 56
Db 293 CATTCACTGAAGGGATGAAGC-----AGT 319
QY 57 GluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrcysVal---AspAsnAsn 75
Db 320 GACTTTATTACAAAGAACAGAAATTTGGTGGAGCTCAGTATTCTGTACACAGAGAAAGA 379
QY 76 GlnGlyGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGlu 95
Db 380 GAAGAAATTCCTGGACGACGAGCTCGACAGCTCTCTCTGATGGCCACAGATTCAGAG 439
QY 96 LysSerArgThrTyrrValAlaArgAsnGlyGluProGluProThrProValValTyrrGly 115
Db 440 TGCAGCAGG----- 448
QY 116 GluLysGluProSerLysGlyAspProAsnThrGluGluIleAtgGlnSerAspGluVal 135
Db 448 ----- 448
QY 136 GlyAspArgAspHisArgArgProGlnGluLysLysLysAlaLysGlyLeuGlyysGlu 155
Db 449 -----AACAAGAGAGAGACCTTAGGAAAAAGAA 475
QY 156 IleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaAla 175
Db 476 GTTTTATTACTGATGACGCGCTTAACACCTTTCAACCCAGAGAGAGAGCTGCAGCT 535
QY 176 LeuCysLysLysTyrrAlaGluLeuGluGluHisArgAsnSerGlnLysGlnMetLys 195
Db 536 CTCTGTGAAGAAATATCTGCTATCTCTGGAAGAAACAGAGAAATGTTTCAGAAACAATGAAG 595
QY 196 LeuLeuGluLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHis 215
Db 596 ATTCTGCAGAGAGAGCAACCCAGATGTTGGAAGAGAAAGTTTCACTTCAGAGTGAACAC 655
QY 216 SerLysAlaValLeuAlaArgSerLysLeuGluSerLysLeuGluSerArgGluLeuGluArgHis 235

```

Db 656 AGCAAGGCCATCTGGCAGAGCAAACTGGAATCTCTTTGAGGGAACTTCAGCGTCAT 715
 QY 236 AsnArgSerLeuLysGluGluValGlnArgAlaArgGluGluGluLysArgLys 255
 Db 716 AATAAGACCTTAAGAGGAGGAATATGAGCAGGACGAGAGGAGGAACACGCTAAA 775
 QY 256 GluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHis 275
 Db 776 GAAGCAACAGCAGCATTTCCAGATACTTAATGAAATCCAAAGCTCAGTTGGAAACACAT 835
 QY 276 AsnGluArgAsnSerLysLeuArgGlnGlnMetGluLeuAlaGluArgLeuLysLys 295
 Db 836 GACATCCCAATGCCAACTCGCAGCAGGAGACATTCAGTGGAGAGAGTGAAGAG 895
 QY 296 LeuIleGluGlnThrGluLeuArgGluHisIleAspLysValPheLysHisLysAsp 315
 Db 896 CTTATTGAGCAGTATGCACTAAGGAGAGAGCATATTGATAAGTATTCAACACAGGAA 955
 QY 316 LeuGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAla 335
 Db 956 TTGCAACACAGCTTGTGGATGCCAACTTCAGCAACACACAGCTGATAAAGAGCT 1015
 QY 336 GluGluArgHisGlnArgGluLysAspPheLeuLeuLysGluLysValGluSerGlnArg 355
 Db 1016 CATGAAAAACATCAGAGAGAGAGAGATTTTATTAAAGAGCAACAGAAATCCAGGCAC 1075
 QY 356 MetCysGluLeuMetLysGlnGlnThrHisLysLysGlnGlnLeuAlaLeuValThr 375
 Db 1076 AATATGACAAATGAACACAGCAGAGTACAACTAAACAGCAGCTTCTCTTTATAG 1135
 QY 376 GluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPhe 395
 Db 1136 GATAAATTTGAAGATTCAGACTATGCGCAAAAGCAATGAATTTTACACCTTC 1195
 QY 396 LysGlnGluMetGluLysMetThrLysLysLysLysLysGluLysGluThrThrMet 415
 Db 1196 AGGCAAGGAAATGGAAGAGATGACAAAGAAATTAATAAATGGAAGAAACAAATAA 1255
 QY 416 TyrArgSerArgTrpGluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThr 435
 Db 1256 TGGCGTACCAATGGGAAACAAATAATAAAGCACTTCTGAGATGCGCGGAGAGAAACT 1315
 QY 436 ValArgAspLysGluLeuGluGluLysLysLysLysLysGlnArgLeuGluLysLys 455
 Db 1316 GTCCGTGATAAGAGTACAGCTTTTCAATATTAATGGAAGCTTAGAAGAGCTGTGC 1375
 QY 456 ArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgValGlnAspLeu 472
 Db 1376 AGGGCTCTTCAGACAGAGAAATGAGCTCAACGAGAGAGTCAAGTCTCCTG 1426

RESULT 4
 CA328534 798 bp mRNA linear EST 09-JUL-2003
 LOCUS UI-M-FY0-ccx-h-10-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone
 DEFINITION IMAGE:6825755 5', mRNA sequence.

ACCESSION CA328534
 VERSION CA328534.1 GI:24546632
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 798)
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: ccapbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 cDNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: PYX-5.
 Location/Qualifiers

FEATURES

source

1. 798
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="CS7BL/6"
 /db_xref="taxon:10090"
 /clone="IMAG:6825755"
 /tissue_type="whole brain"
 /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
 /lab_host="DHI0B (Ti phase resistant)"
 /clone_lib="NIH_BMAP_FY0"
 /notes="Organ: Brain; Vector: PYX-Asc; Site 1: Ecor I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with Ecor I adaptor, digested with NotI and then cloned directionally into PYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCAGACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemlin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:
 Pred. No.: 1-86e-74 Length: 798
 Score: 1279.00 Matches: 258
 Percent Similarity: 98.86% Conservative: 2
 Best Local Similarity: 98.10% Mismatches: 3
 Query Match: 47.34% Indels: 1
 DB: 14 Gaps: 0
 US-10-023-523-8 (1-530) x CA328534 (1-798)

QY 120 SerLysGlyAspProAsnThrGluLysLeuArgGlnSerAspGluValGlyAspArg 139
 Db 10 TCTAAGGAGAGGCTTGGACAGAGGAGATCCGAGCGATGATGAAGTTGGAGACCGAGAC 69
 QY 140 HisArgArgProGlnGluLysLysLysAlaLysGlyLeuGlyLysGluLeuThrLeu 159
 Db 70 CATCGAGGCGCACAGAGAGAAAGCAAGGGTCTAGGAGAGGAGATCACTCTGCTG 129
 QY 160 MetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLys 179
 Db 130 ATGCAGACACTGACACCGCTGAGTACCCAGAGAGAGAGCTGGCTGCCTGTGCAGAGAG 189
 QY 180 TyrAlaGluLeuGluGlnHisArgAsnSerGlnLysGlnMetLysLeuGlnLys 199
 Db 190 TATGCTGAGCTGCTGGAAGAGCATCGGAATCGCAGAGAGAGATGAGCTCTCTCAGAGAG 249
 QY 200 LysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaVal 219
 Db 250 AACAGAGCCAGCTCTGTGAGAGAGGACCATCTGCGAGGGGAGACACAGAGAGGCTGTC 309
 QY 220 LeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeu 239
 Db 310 CTGCGCCGAGCAAGCTTGAAGAGTCTGTGCCGGAGCTGCAACCGCACACCGTCCCTG 369
 QY 240 LysGluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSer 259
 Db 370 AAGAGAGAGGCGTGCAGCGAGCCCGTGGAGAGAGAGAGAGCCCAAGAGAGTGCATCA 429
 QY 260 HisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsn 279

DEFINITION UI-M-HB0-ckc-1-18-0-UI.r1 NIH BMAP_HB0 Mus musculus cDNA clone
IMAGE:30615737 5', mRNA sequence.

ACCESSION CF735205

VERSION CF735205.1 GI:37631541

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 770)

AUTHORS Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE NIH-MGC http://mgc.nci.nih.gov/

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue procurement: Dr. James Lin University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1..770

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:30615737"

/tissue_type="whole eye"

/dev_stages="embryo 12.5,13.5,14.5 dpc"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NIH BMAP_HB0"

/note="Organ: Eye; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is TATGTAGAGT. This library was created for the University
of Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH)."FEATURES
source

ORIGIN

Alignment Scores:

Pred. No.:	1.72e-73	Length:	770
Score:	1264.00	Matches:	253
Percent Similarity:	99.22%	Conservative:	1
Best Local Similarity:	98.83%	Mismatches:	2
Query Match:	46.78%	Indels:	0
DB:	14	Gaps:	0

US-10-023-523-8. (1-530) x CF735205 (1-770)

Qy 147 LysLysAlaLysGlyLeuGlyLysGluLeuThrLeuLeuMetGlnThrLeuAanThrLeu 166

Db 2 AAGAAAGCCAAAGGCTAGGAGAGAGATCACTCTGCTGATGAGACACTGAACAGCTG 61

Qy 167 SerThrProGluGluLysLeuAlaLeuCysLysLysTyrAlaGluLeuLeuGlu 186

Db 62 AGTACCCAGAGAGAGAGCTGGCTGCATGTGCAAGAGATGCTGAGCTGCTGGAAGAG 121

Qy 187 HisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysGlnSerGlnLeuValGln 206

Db 122 CATCGGAACTCGCAAGAGCAGATCAAGCTCTCTGAGAAAGAGCAGACGCTCGTGAC 181

Qy 207 GluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGlu 226

Db 182 GAGAAGGACCAATCTGGAGGGGACACACCAAGGCTGCTCTGCCCGAAGCAAGCTTGAG 241

Qy 227 SerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGluGlyValGlnArg 246

Db 242 AGTCTGTGCGGAGGCTGCAAGCGCACACACCGGTCCCTGAAGGAAAGAGCGCTGAGGGA 301

Qy 247 AlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsn 266

Db 302 GCCGTGAGAGAGAGAGAGAGCGCAAGAGTGCATTCACACTTCAGGTGACACTGAT 361

Qy 267 AspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsn 286

Db 362 GACATTCAGCTGCAGATGGAACAGCATAACGAGCAGCAAACTCCCAAGTCGCCAGGAGAAT 421

Qy 287 MetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTyrGluLeuArgGluGluHis 306

Db 422 ATGGAGCTAGCCGAGGCTCAAGAGTTGATCGAGCAATAGCAGCTTCGTGAGGAGCAT 481

Qy 307 IleAspLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGln 326

Db 482 ATCGCAAAAGTCTTCAACATCAAGCAGCTGCGAGCAGCAGTGTGCGACGCCAAGCTCCAG 541

Qy 327 GlnAlaGlnGluMetLeuLysGluAlaGluGluArgHisGlnArgGluLysAspPheLeu 346

Db 542 CAGGCCCGAGGATGCTGAAGGAGCAGAGGAGCGGCGCACCGAGGAGAGGATTTCTC 601

Qy 347 LeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHis 366

Db 602 CTGAGGAGCGGTGGAGTCCCAGAGGATGTGCGAGCTGATGAACGACGAGAGACCCAC 661

Qy 367 LeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSer 386

Db 662 CTAAGACAGCAGCTCGCCCTGTACACGAGAGAGTTTGAGGAGTTCCAGAACACACTTTC 721

Qy 387 LysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMet 402

Db 722 ANAGCAGTGAAGTGTTCACACGTTCAACAGGAGATGGNAGATG 769

RESULT 7

BG827157

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

932 bp mRNA linear EST 22-MAY-2001
602751143F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4903690 5',
mRNA sequence.

BG827157

BG827157.1 GI:14174744

EST.

Homo sapiens

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 932)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: L10M1802 row: f column: 11

High quality sequence start: 6

High quality sequence stop: 775.

Location/Qualifiers

1..932

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4903690"
/tissue type="rhombomysorcoma"
/lab host="DH10B (phage-resistant)"
/clone lib="NH MGC 17"
/note="Organ: muscle; Vector: pOTB7; Site 1: EcoRI;
Site 2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGACAGG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Alignment Scores:			
Pred. No.:	1.68e-72	Length:	932
Score:	1250.50	Matches:	274
Percent Similarity:	95.82%	Conservative:	1
Best Local Similarity:	95.47%	Mismatches:	6
Query Match:	46.28%	Indels:	9
DB:	12	Gaps:	0
US-10-023-523-8 (1-530) x BG827157 (1-932)			
QY	241	GlUGluGlyValGlnA-gAlaArgGlUGluGluGluLysArgLysGluValThrSerHis	260
DB	6	GAAGAAGTGTGCAGCGCGGCCCGGAGGAGGAGCGAAGCGAGGTGACCTCGCAC	65
QY	261	PheGlnValThrLeuAsnAspIleGlnLeuGluMetGluGlnHisAsnGluArgAsnSer	280
DB	66	TTCCAGGTGACACATGAATGACATTACGCTCCAGATGGAACAGACAATGAGCGCAACTCC	125
QY	281	LysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTyr	300
DB	126	AACTGCCCAAGAGACATGAGCTGGCTGAGAGGCTCAAGAGCTGATTGAGCAGTAT	185
QY	301	GluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeu	320
DB	186	GAGCTGCGGAGGAGCATATCGACAAAGTCTTCAAAACACAAGGACCTCAACAGCAGCTG	245
QY	321	ValAspAlaLysLeu-GlnGlnAlaGlnGluMetLeuLysGluAlaGluGluArgHisGln	340
DB	246	GTGGATGCCAAGCTTCAGCAGGCCAGGAGATGCTAAAGAGGCGAGAAAGCGGCACCA	305
QY	340	nArgGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMe	360
DB	306	CGCGGAGAGGATTTCTCTCTGAAAGAGGCGAGTAGAGTCCAGAGGATGTGTGASCTGAT	365
QY	360	tlLysGlnGlnThrHisLysLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGln	380
DB	366	GAAGCAGCAAGAGACCCACCTGAAGCAACAGCTTGCCTTATACACAGAAAGTTTGAGGA	425
QY	380	pPheGlnAsnThrLeuSerLysSerGluValPheThrThrPheLysGlnGluMetGln	400
DB	426	GTTCCAGAACACACTTTCACAAAGACAGCGAGGTATTCACCAATTCAACAGCAGGATGGA	485
QY	400	uLysMetThrLysLysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArg-T	420
DB	486	AAAGATGACTAAGAAGATCAAGAAGCTCGAGAAAGAAAACACCATGTACCGTCCCGGTT	545
QY	420	rpGluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAsp-Lys	439
DB	546	GGGAGAGCAGCAACAAAGGCCCTGCTTGAGATGGCTGAGGAGAAAACAGTCCCGGACTAAA	605
QY	440	GluLeuGluGlyLeuGln-ValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGln	459
DB	606	GAACTGGAGGGCTGCAGGGTAAAAATCCAAACGGCTGGAGAAAGCTGTGCCGGGCACCTGCA	665
QY	459	nThrGluArgAsnAspLeuAsn-LysArgValGlnAspLeuSerAlaGlyGlnGlnLys	479
DB	666	GACAGAGCGCAATGATACCTGAACCAAGAGGGGTACAGGACCTGAGTGTGTGTGCCAGGGCT	725
QY	479	erLeuThrAspSerGlyProGluArgArgProGluGlyProGluValAlaGlnAlaProSer-	498

Db	726	C-CTCACTGACAGTGG-CTTGAGAGGAGCA-GAGGGGCGCTGGGGCTCAAGCACC	CCCAGG	782
Qy	499	SexProArgValThrGluAlaProCysTyrProGlyValaProSerThrGluAlaSerGly	518	
Db	783	TCCCCCAGGGTCAcAGAGGGCGCTTGTCACCCAGGAGCACCGAGCACAAGATCA	842	
Qy	519	GlnThrGly	521	
Db	843	AGACTGGGC	851	

RESULT 8
CA320886

CA320886 806 bp mRNA linear EST 09-JUL-2003
UI-M-690-cdb-h-16-0-UI.r1 NIH-BMAP_FW0 Mus musculus cDNA clone
IMAGE:6817313 5', mRNA sequence.
CA320886
CA320886.1 GI:24538984
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 806)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-t@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

FEATURES
source

```

Seq primer: pyx-5.
Location/Qualifiers
1. 806
/organism="Mus musculus"
/mol_type="mrna"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6817313"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dp"
/lab_host="DH10B (Ti phase resistant)"
/clone_lib="NIH BMAP FWO"
/note="Organ: Brain; Vector: pyx- Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according to Bontaïdo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pyx-Asc vector. The library tag sequence located between the Not I site and the polyA tail is ACGGACACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

```

ORIGIN

Alignment Scores:	
Pred. No.:	4, 06e-72
Score:	1243.50
Percent Similarity:	97.01%
Best Local Similarity:	96.27%
Query Match:	46.02%
Length:	806
Matches:	258
Conservative:	7
Mismatches:	7
Indels:	3


```

493 GAGGAGTACAGAGGCTGTCGAGGAGGAGAGAAACGCAAGAGTAAGTAACTCCCTTC 434
262 GlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLys 281
433 CAAAGTACAGCTCAATGACATCCAGTACAGATGGAGGAGCAGCAATGACGTAATGCTAAA 374
282 LeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTyrGlu 301
373 CTCGCGTCAAGAGAAATGTTGAGCTTCTGATCGACTTAAGAGAGCTTATCGAGCAATGAG 314
302 LeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnGlnLeuVal 321
313 CTAAGAGAGAGACATTCACAAAGTTCACAAATAGGATCTTCACAGCACTGGTG 254
322 AspAlaLysLeuGlnGlnAlaGlnMetLeuLysGluAlaGluGluArgHisGlnArg 341
253 GATGCGAAGCTTCAGCAAGCAGCAAGAAATGCTGAAGAAGTGGAGAGCGCTCCACGCG 194
342 GluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLys 361
193 GAAAGAGAGTTCCTGTTGAAGAGGAGCAGTCCGAATCCCGAGCGCATGTGTGAACATAAGAAA 134
362 GlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPhe 381
133 CAGCAAGAGACCCACCTCAACACAGTACGACTGTATACAGAGAAAGTTTGAGGAGTTT 74
382 GlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLys 401
73 CAAACACCTTCTCTAAACACAGTGGAGTTTCACCACTTTTAA-CAAGAAATGGAAAAA 15
402 MetThrLysLys 405
14 AAAAAAAAAAAAA 3

```

```

RESULT 10
CB247992
LOCUS
DEFINITION
  CB247992 779 bp mRNA linear EST 09-JUL-2003
  UI-N-F10-ceb-c-22-0-UI-r1 NIH_BMAP_F10 Mus musculus cDNA clone
  IMAGE:6837167 5', mRNA sequence.
ACCESSION
  CB247992
VERSION
  CB247992.1 GI:28369636
SOURCE
  Mus musculus (house mouse)
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 779)
  NIH-MGC http://mgi.nci.nih.gov/
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  Contact: Robert Strausberg, Ph.D.
  Email: cgaabs-remail.nih.gov
  Tissue Procurement: Dr. Jim Lin, University of Iowa
  cDNA library preparation: Dr. M. Bento Soares, University of Iowa
  DNA library arrayed by: Dr. M. Bento Soares, University of Iowa
  DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
  Clone Distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  http://image.llnl.gov
  This clone was contributed by the Brain Molecular Anatomy Project
  (BMAP)
  Seq primer: pYX-5,
  Location/Qualifiers
    1..779
      /organism="Mus musculus"
      /mol_type="mRNA"
      /strain="C57BL/6"
      /db_xref="taxon:10090"
      /clone="IMAGE:6837167"
      /tissue_type="whole brain"
      /dev_stage="embryo 12.5dpc"
      /lab_host="DH10B (T1 phage resistant)"

```

FEATURES
source

ORIGIN

```

Alignment Scores:
Pred. No.: 1,52e-71 Length: 779
Score: 1234.50 Matches: 251
Percent Similarity: 97.31% Conservative: 2
Best Local Similarity: 96.54% Mismatches: 6
Query Match: 45.69% Indels: 2
DB: 14 Gaps: 1

US-10-023-523-8 (1-530) x CB247992 (1-779)
QY 103 ArgAsnGlyGluProGluPro---ThrProValValTyrGlyGluLysGluProSerLys 121
Db 1 AGGAATGGGAGGCTCGAACCCAGGCAATCCAGTCGTCAACGGCGAGAGGAGACCTCTAAG 60
QY 122 GlyAspProLeuThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHisArg 141
Db 61 GGAGAGCTCGAAGCAGAGGAGATCCGAGCGAGTGAAGTGGAGACCGAGACCATCGG 120
QY 142 ArgProGlnGluLysLysLysAlaLysGlyLysGluLysGluLysLeuLeuMetGln 161
Db 121 AGGCCACAGAGAGAGAGAAAGCCAGGGTCTAGGCGAGGAGATCACTCTGTGTGTCAG 180
QY 162 ThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuLysLysLysTyrAla 181
Db 181 ACCTGAACACGCTGAGTACCCAGAGGAGAGCTGGCTGCACATGTGCAAGAGTATGCT 240
QY 182 GluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGln 201
Db 241 GAGCTGCTGGAAGAGCATCGAATCTCGAGAACACACATGAACTCTCTCGAGAGAGCAG 300
QY 202 SerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAla 221
Db 301 AGCCAGCTCTGTGAGGAGAGAGGACCATCTCGAGGGGGAACACAGCAAGGCTGTCTTGGCC 360
QY 222 ArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGlu 241
Db 361 CGAAGCAAGCTTCAGAGTCTGTCCGGAGCTGCAACGGCACACACCGGTCCCTGAAGGAA 420
QY 242 GluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPhe 261
Db 421 GAAGGCTGTGAGGAGGAGCCCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
QY 262 GlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLys 281
Db 481 CAGGTGACATCTGAATGACATTCAGTTCAGATGGAACACAGCATAAACGAGCAAACTCCAAG 540
QY 282 LeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLeuIleGluGlnTyrGlu 301
Db 541 CTGCGCCAGAGAGAAATATGGAGCTAGCGAGAGCTCAAGAGTGTATCGAGCAATACGAG 600
QY 302 LeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeuVal 321
Db 601 CTTCTGTGAGGAGCATATCGCAAAAGTCTTCAACATAAGGACCTCGAGCAGCAGCAGTAGTG 660
QY 322 AspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGlnArg 341

```

/clone_lib="NIH_BMAP_F10"
/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is CAGCCACGAC. This library was created for the PolyA tail
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
Program coordinator."

661	GAGCCAGCTCCAGCAGGCCAGGAGATCTGAGGAGGCAGGAGCGGCACGAGGA	720
Db		
342	GIuLySAspPheLeuLeuLySGluAlaValGIuSerGIuArgMetCysGIuLeuMetIys	361
Qy		
721	GAGAGGAGTTTCTCTGAAGAGCGGTGGATC-CAGAGGATGTGCAGGTGATGAAG	779
Db		

RESULT 11

Accession	LOCUS	Size	Source	Accession	EST
AL961454	LOCUS	945 bp	mRNA	linear	EST 04-DEC-2003
AL961454	DEFINITION	AL961454 XGC-gastrula Silurana tropicalis cDNA clone TGas127p15', mRNA sequence.			

ACCESSION	AL961454
VERSION	AL961454.2
KEYWORDS	GI:38703803
SOURCE	EST.
ORGANISM	Silurana tropicalis (western clawed frog)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Silurana.
1 (bases 1 to 945)

REFERENCE

Croning, M.D.R., Ashurst, J.L., Taylor, R., Zorn, A.M. and Rogers, J.
 Sanger Xenopus tropicalis EST project 2001 (11_2003)
 Unpublished (2003)
 On Nov 27, 2002 this sequence version replaced gi:25785049.

Qy	103	A-rAsnGlyGluProGlu---ProThrProValValTyrGlyGluLysGluProSerLys	121
Db	93	AAAATGGTGACTACTAGTACCAAGAGGTTAATGGTCAGAAAGAGGAAATCGCG	152
Qy	122	GlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspAspHisAsxg	141
Db	153	GTTTTAGAAGACAGTAGAACCATCAACGAGACT-----GAAAAGATCAGAA	200
Qy	142	ArgProGlnGluLysGlyLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGln	161
Db	201	CGAATGCCAAGCAAAAAAAGCGAAGGGTCGGGAAGGAGATCACATTATTAAATGCAG	260
Qy	162	ThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaAlaLeuCyLysLysTyrAla	181
Db	261	ACTCTGAACACACTCAGCACCCAGAGAAAGCTGACTGCCCTGTGTAAAGAAATATGCA	320
Qy	182	GluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGln	201
Db	321	GAACCTGTTGGAGGACACAGGACATCTCAGAAGCAAAATAGGATACTGCAAAAACACAG	380
Qy	202	SerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAla	221
Db	381	ACCCAGCTTATTCAAGAAAGGATCTACTTCGTAAATGAGCACAGCAAGGCTATCTCGCA	440
Qy	222	ArgSerLysLeuGluSerLeuCyAsxGluLeuGlnArgHisAsnArgSerLeuLysGlu	241
Db	441	AGAGCAAGCTCGAAAGTTGTGCGAGAGCTACAGAGACACAAACCGTACCTTAAGGAG	500
Qy	242	GluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPhe	261
Db	501	GAAGAGGTACAGAGGGCTCGTGAGGAGGAAGAGAAACGAAAGAAAGTAACCTCCCATTC	560
Qy	262	GlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLys	281
Db	561	CAAGTGACGCTCAATGACATCCAGTCACAGATGGAGCAGCACAACTGAACGTAACTGATA	620
Qy	282	LeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLeuIleGluGlnTyrGlu	301
Db	621	CTGCGCTCAAGAGAATGTGAGCTTGCTGATCGACTTAAGAAAGCTTATCAGCAAAATGAG	680
Qy	302	LeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeuVal	321
Db	681	CTAAGAGAAGAGCACATTGACAAAGTTTCAACATAAGGATCTTCAACAGCAACTGGTG	740
Qy	322	AspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluLysGlnArg	341
Db	741	GAITCAAAAGCTTCAGCAAGCACAGAATAATGTTGAAGAAGTGAAGAGCGTCACCAAGCG	800
Qy	342	GluLysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLys	361
Db	801	GAAAGAGGAGTTCTGTGTTGAAGAGGCGAGTCCGATCCAGCGCATGTGTGCAACTAATGAA	860
Qy	362	GlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPhe	381
Db	861	CAGCAAGAGACCCACCTCAACACACAGTTAGCTGTATACAGAGAAAGTTTGAGGAGTTT	920
Qy	382	GlnAsnThrLeuSerLysSerSer	389
Db	921	CAAAACACCTTGTCTAAAGCAAT	944

[illegible]

AUTHORS TITLE JOURNAL COMMENT

NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
DNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5.

FEATURES source

Location/Qualifiers
1..725
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5697868"
/tissue_type="whole brain"
/dev_stage="embryo 12.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP F10"
/note="Organ: Brain; Vector: pYX-Asc; Site 1: Ecor I;
Site 2: Not 1; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dr
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with Ecor I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is CAGCCACGAC. This library was created for the University
Iowa Brain Anatomy Project (BMAP): Gene Discovery in the
Developing Mouse Nervous System, supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

ORIGIN

Alignment Scores:
Pred. No.: 6.63e-70 Length: 725
Score: 1209.00 Matches: 240
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.59% Mismatches: 0
Query Match: 44.74% Indels: 0
DB: 13 Gaps: 0

US-10-023-523-8 (1-530) x BU701958 (1-725)

QY 136 GlyAspAcqAspHisArgArgProGlnGlyLysLysAlaLysGlyLeuGlyLysGlu 155
DB 1 GGAGACCGAGACATCGAGGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
QY 156 IleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGlnGlyLysLeuAla 175
DB 61 ATCACTCTCTGATGACACACCTGACCTGAGTACCCAGAGAGAGAGAGAGAGAG 120
QY 176 LeuCysLysLysThrAlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLys 195
DB 121 CTGTGCAAGAGATGCTGAGTCTGAGAGAGATCGAATCTCGAGAGAGAGAGAGAG 180
QY 196 LeuLeuGlnLysLysGlnSerGlnLeuValGlnGlnLysAspHisLeuArgGlyGluHis 215
DB 181 CTCCTGCAGAC 240
QY 216 SerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHis 235
DB 241 AGCAAGGCTGCTCTGCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300

QY 236 AsnArgSerLeuLysGlnGlnGlyValGlnArgAlaArgGlnGlnGlnGlnLysArgLys 255
DB 301 AACCGGTCCCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 256 GluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuMetGlnGlnHis 275
DB 361 GAATGACTTCCACATTCAGGTCGACACTGAATGACATTCAGTCAGATGAAGAGAT 420
QY 276 AsnGluArgAsnSerLysLeuArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 295
DB 421 AACGAGCGAAATCCCAAGCTCGCCAGGAGATATGGAGCTAGCCGAGAGGCTCAAG 480
QY 296 LeuIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 315
DB 481 TTGATCGAGCAATACGAGCTTCGTGAGGAGCATATCGAACAAGTCTTCAACATAG 540
QY 316 LeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnGlnGlnGlnGlnGlnGln 335
DB 541 CTGCAGCAGCAGCTAGTGGAGCGCCAAAGCTCCAGCAGCCCGAGAGATGCTGAAG 600
QY 336 GluGluArgHisGlnArgGlnLysAspPheLeuLysGlnGlnGlnGlnGlnGln 355
DB 601 GAGGAGCGGACACGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 356 MetCysGlnLeuMetLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 375
DB 661 ATGTGCGAGCTGATGAAGCAGCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 376 Glu 376
DB 721 GAG 723

RESULT 13

BQ769910 735 bp mRNA linear EST 26-JUL-2002
UI-M-F10-byt-1-21-0-UI-r1 NIH BMAP_F10 Mus musculus cDNA clone
IMAGE:5702108 5', mRNA sequence.

ACCESSION BQ769910
VERSION BQ769910.1 GI:21978384

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 735)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1..735

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:5702108"

/tissue_type="whole brain"

/dev_stage="embryo 12.5dpc"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NIH BMAP F10"

/note="Organ: Brain; Vector: pYX-Asc; Site 1: Ecor I;

Site 2: Not 1; the library was constructed according

Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CAGCACGAC. This library was created for the University of Iowa Brain Anatomy Project (BMAP): "Gene Discovery in the Developing Mouse Nervous System", supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:
Pred. No.: 9,138-70 Length: 735
Score: 1207.00 Matches: 239
Percent Similarity: 100.00% Conservativity: 2
Best Local Similarity: 99.17% Mismatches: 0
Query Match: 44.67% Indels: 0
DB: 13 Gaps: 0

US-10-023-523-8 (1-530) x BQ769910 (1-735)

QY 135 ValGlyAspArgAspHisArgArgProGlnGluLysHisLysAlaLysGlyLeuGlyLys 154
DB 11 ATTGGAGACCGAGACCATCGAGGCGCACAGGAGAGAAAGCAAGGCTTAGGGAAG 70
QY 155 GluIleThrLeuMetGlnThrLeuSerThrLeuSerThrProGluGluLysLeuAla 174
DB 71 GAGATCCTCTCTGATGACACATGAAACCGCTGATCCCGAGGAGAGAGCTGGCT 130
QY 175 AlaLeuCysLysLysTyrrAlaGluLeuLeuGluGluHisArgSerGlnLysGlnMet 194
DB 131 GCACCTGTGCAAGAGTATGCTGAGCTGTGGAAGAGCATCGGAATCGCAGAGCAGATG 190
QY 195 LysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGlu 214
DB 191 AAGCTCTCTGAG 250
QY 215 HisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArg 234
DB 251 CACAGCAAGGCTGTCTCTGCGCGGAGCAAGCTTGAGAGTCTGTGCGCGGAGCTGCAACGG 310
QY 235 HisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGluGluGluLysArg 254
DB 311 CACAACCGGCTCTCTGAG 370
QY 255 LysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGln 274
DB 371 AAAGAAGTGACCTTACATCTCCAGTGACATCGAATGACATTCAGCTGCAGATGGAGACAG 430
QY 275 HisAsnGluArgAsnSerLysLeuArgGlnGlnLeuMetGluLeuAlaGluArgLeuLys 294
DB 431 CATACAGAGCGCAACTCCAGCTGCGCAGAGAGATATGAGCTAGCCGAGAGAGCTCAAG 490
QY 295 LysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAspLysValPheLysHisLys 314
DB 491 AAGTTGATCGACATACGAGCTCTGTCGAGGAGCATATCGACAAGTCTTCAACACATAAG 550
QY 315 AspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGlu 334
DB 551 GACCTGCGAGCAGCAGCTAGTGAGCGCCAGCTCCAGCAGGCCCGCAGAGATGCTGAAGAG 610
QY 335 AlaGluGluArgHisGlnArgGluLysAspPheLeuLeuLysGluAlaValGluSerGln 354
DB 611 GCAGAGGAGCGGCACACAGCAGAGAGAGAGTCTCTCTGGAAGAGCGGTGAGTCCCGAG 670
QY 355 ArgMetCysGluLeuMetLysGlnGlnGlnThrHisLeuLysGlnGlnLeuAlaLeuTyr 374
DB 671 AGGATGTGCGAGCTGATGAGACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 730

QY 375 Thr 375
DB 731 Arg 733

RESULT 14
CF995807

LOCUS

CF995807 759 bp mRNA linear EST 25-NOV-2003
AGENCY: NIH MGC 221 Homo sapiens cDNA clone
IMAGE: 30708497 5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 759)

NIH-MGC http://mgs.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics / NIH

National Cancer Institute

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgsaps-remail.nih.gov

Tissue Procurement: James Martin, University of Iowa

cDNA Library Preparation: M. Bento Soares, University of Iowa

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

http://image.llni.gov

Plate: NDAMI074 row: m column: 18

High quality sequence stop: 631.

Location/Qualifiers

1..759

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:30708497"

/lab_host="DH10B T0A"

/clone_lib="NIH MGC 221"

/note="Organ: mixed; Vector: pYX-Asc; Site 1: EcoRI;

Site 2: NotI; Library is oligo-dT primed and directionally

cloned. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT

primer containing a Not I site. Double strand cDNA was

size selected according to mRNA size fraction, ligated with

EcoR I adaptor, digested with Not I and then cloned

directionally into pYX-Asc vector. Average insert size

4-5kb. Adaptors 5' (AATTCGACAGG) 3' and 5' d

(CTCTGTGCG) 3'. Linker sequence - GCGCGCTGTGAGGCC T18.

Sequencing primers 3' end: T3 promoter primer 5' d

(ATTAACCTCTACTATAGG) 3'. 5' End: T7 promoter primer 5' d

(TAATACGACTCTATAGG) 3'. Library was constructed in the

laboratory of M. Bento Soares. Note: this is a NIH_MGC

Library"

ORIGIN

Alignment Scores:

Pred. No.: 1,298-69 Length: 759

Score: 1205.00 Matches: 242

Percent Similarity: 98.39% Conservativity: 2

Best Local Similarity: 97.58% Mismatches: 4

Query Match: 44.60% Indels: 0

DB: 14 Gaps: 0

US-10-023-523-8 (1-530) x CF995807 (1-759)

QY 188 ArgAsnSerGlnLysGlnMetLysLeuGlnLysGlnSerGlnLeuValGlnGlu 207

DB 9 CGGATTCACAGAGAGCAGATGATGAGCTCTCTACAGAGAGAGAGAGAGAGAGAG 68

QY 208 LysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSer 227
 Db 69 AAGGACCACTCGCGGTGACACAGCAAGCGCTCTGGCCCGCAGCAGCTTGAGAGC 128
 QY 228 LeuCyArgGluLeuGlnArgHisAsnArgSerLeuLysGluGlyValGlnArgAla 247
 Db 129 CTATCCGTGAGCTGACGCGCACACACCGCTCCCTCAAGGAAGAAGGTGTGACGGGCC 188
 QY 248 ArgGluGluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAsp 267
 Db 189 CGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 248
 QY 268 IleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMet 287
 Db 249 ATTCAAGTGCAGATGGAACACAGCAATGAGCGCAACTCCAGCTGCGCAGAGAACATG 308
 QY 288 GluLeuAlaGluArgLeuLysLysLeuIleGlnThrGluLeuArgGluHisIle 307
 Db 309 GAGCTGGCTGAGGCTCAAGAAGCTGATTGAGCAGTATGAGCTGCGCGAGGAGCATATC 368
 QY 308 AspLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAlaLysLeuGlnGln 327
 Db 369 GACAAAGTCTTCAACACAGAGCACTACACAGCAGCTGGTGGATGCCAGCTCCAGCAG 428
 QY 328 AlaGlnGluMetLeuLysGluAlaGluArgHisGlnArgGluLysAspPheLeuLeu 347
 Db 429 GCCCAGGAGATGCTAAAGAGGCGCAGAGAGCGGACCAAGCGGAGAGAGGATTTCTCCTG 488
 QY 348 LysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLeu 367
 Db 489 AAGAGGCGAGTAGAGTCCAGAGAGTGTGAGCTGATGAAGCAGCAGAGACCCACCTG 548
 QY 368 LysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLysSerLys 387
 Db 549 AAGCAACAGCTTGCCTATACACAGAGAGTTTCAGGAGTTCCAGACACACATTTCCAAA 608
 QY 388 SerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysIleLys 407
 Db 609 AGCAGCGAGGTATTACACACATNTCAGCAGAGATGGAAGAAAGNATGACTAGAAGATCCAG 668
 QY 408 LysLeuGluLysGluThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeu 427
 Db 669 AAGCTGGAGAAAGAAACACCATGTACCGTCCCGTGGGAGAGCAGCAGACAGGCCCTG 728
 QY 428 LeuGluMetAlaGluGluLysThr 435
 Db 729 CTTGAGATGGCTGAAGAAAAACA 752

RESULT 15
 BG761234
 LOCUS 602718583P1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4858388 5',
 DEFINITION mRNA sequence.
 ACCESSION BG761234
 VERSION BG761234.1 GI:14071887
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 817)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC/DCTP/DTF
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: LLCM1712 row: f column: 21
 High quality sequence stop: 766.
 Location/Qualifiers
 source 1..817
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4858388"
 /tissue_type="melanotic melanoma, high MDR (cell line)"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 49"
 /notes="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCACGAG(G). Size-selected ~500bp for average insert size
 1.8kb. Library constructed by Ling Hong in the laboratory
 of Gerald M. Rubin (University of California, Berkeley)
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC
 Library. |"
 ORIGIN
 Alignment Scores: Length: 817
 Pred. No.: 1.9e-69 Matches: 246
 Score: 1203.00 Conservative: 0
 Percent Similarity: 96.85% Mismatches: 4
 Best Local Similarity: 96.85% Indels: 4
 Query Match: 44.52% Gaps: 0
 DB: 12
 US-10-023-523-8 (1-530) x BG761234 (1-817)
 QY 1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyValaGlnGluArgProSerGln 20
 Db 63 AAAAGCAGCCAGCAGCAACCGGAAGCAGGAGCCCGAGGAGCCAGGAGCGGCCAGCCAG 122
 QY 21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40
 Db 123 GCGGCTCTCTGAGTAGAAGCAGAGGTCCCGCAGCAGCAGGCTCTCTCGAAGCGGAG 182
 QY 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 60
 Db 183 GGGGCTCAGCCAGCAGACGCTCAGTCTGGGGCCCTTCGTGTGTCTCTGAGAGCTGAGC 242
 QY 61 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGly 80
 Db 243 CGCCAACTGGAAGACATACCTAGACATACATCTGTGGACAATAAACCAGGGGGGCCCGGC 302
 QY 81 GluAspGlyValaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
 Db 303 GAGATGGGGCAGCAGGCTGAGCGGCTGAACCCGAGATGCGAGAGAGTCCCGGACCTAT 362
 QY 101 ValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGluLysGluProSer 120
 Db 363 GTGGCAAGGAATGGGAGCCTGAACCACTCCAGTAGTCAATGGAGAGAAGAACCTCC 422
 QY 121 LysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHis 140
 Db 423 AAGGGGATCCAAACACAGAGAGATCCGCGAGAGTGCAGGCTCGAGACCGAGACCAT 482
 QY 141 ArgArgProGlnGluLysLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMet 160
 Db 483 CGAAGCCACAGGAG 542
 QY 161 GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysTyr 180
 Db 543 CAGACATTTGAATCTCTGATGATCCCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAT 602
 QY 181 AlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLys 200
 Db 603 GCTGAAGTCTGAG 662
 QY 201 GlnSerGlnLeu-ValGlnGluLys-AspHisLeuArgGlyGluHisSerLysAlaVal- 219

Db 663 CAGACCCAGCTGGTGCAGAGAGAGGACACCTGGCGGTGACACAGCAGCGCTCC 722
 QY 220 leuAlaArg-SerIysLeuGluSerIysCysArgGluLeuGlnArgHisAsnArgSerLe 239
 Db 723 TTGGCCCGCCAGCAAGCTTGAGAGTCTATGCCGTGAGCTGCGGCGGCAACAACATCCCT 782
 QY 239 uIysGluGluGlyValGlnArgAlaArgGluGlu 250
 Db 783 CAAGGAGAGAGTGTGCAGCGGCGGCCGAGGAA 816

AK084639 1608 bp mRNA linear HTC 20-SEP-2003
 Mus musculus 13 days embryo heart cDNA, RIKEN full-length enriched
 library, clone:D330024K18 product:MUSCLE-DERIVED PROTEIN MDP77
 VARIANT 1, full insert sequence.
 AK084639.1 GI:26351202
 AK084639 HTC; CAP trapper.
 Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1. Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 9279253
 10349636

2. Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 2049374
 11042159

3. Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861

4. The RIKEN Genome Exploration Research Group Phase II Team and the
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)
 5. The PANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II team.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6 (bases 1 to 1608)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
 Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
 Katoh, H., Kawai, Y., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
 Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Tagami, M., Tegawa, A., Takanashi, F., Takaku-Akashira, S.,

Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
 Muramatsu, M. and Hayashizaki, Y.
 Direct Submission
 Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/
 URL: http://fantom.gsc.riken.go.jp/
 URL: http://fantom.qualifiers

FEATURES
 source
 1. .1608
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="PANTOM DB:D330024K18"
 /db_xref="MGI:2421329"
 /db_xref="taxon:10090"
 /clone="D330024K18"
 /tissue_type="heart"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="13 days embryo"
 218. _->1606
 /note="unnamed protein product; MUSCLE-DERIVED PROTEIN
 MDP77 VARIANT 1 (SPTR:AAJ33909, evidence: FASTY, 99.1% ID,
 100% length, match=2058)
 putative"
 /codon_start=1
 /protein_id="BAC39238.1"
 /db_xref="GI:26351203"
 /translation="MEINHPDQLSVHEPTPGDSSSLNQNGPKODGRCSTSGOAPE
 QGSLHTEPGADHVAELSESRQEDLIISYGSASPRGKESITSEKEQPPNTEAPNED
 VYQETEDIDREPTAPEAPAAKPFVSNKQKELKGLGKREANLLMQNLKLOA
 PEKDLFPKYTELDETEQKLLKQAOQOTREKQDQSEHNRAVLARSKLE
 SICRELQRNKTKEETLQAREEEREKEITSHFTITDITQITBOOSEERNKLCC
 ENTELAELKLSIIDVELREHLDKFXHRELQOKLAVKLEAEALQMOEAERHRE
 KYLLNQAEWKLOAKLKEQTVLQAQTLYSGRFEFPOSTLTIKSNVEVATFKQEMD
 KTKKMKLEKDTATWKFENCNKALLDMIEKALRAKEYECFVWKIQRLNLCRAL
 QERKELYKKIRE"

Alignment Scores:
 Pred. No.: 1,7e-68 Length: 1608
 Score: 1194.00 Matches: 251
 Percent Similarity: 67.86% Conservative: 72
 Best Local Similarity: 52.73% Mismatches: 103
 Query Match: 44.19% Indels: 50
 DB: 11 Gaps: 6

US-10-023-523-8 (1-530) x AK084639 (1-1608)

QY 18 ProSerGlnAlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArg 37
 Db 263 CCCCCAGGGGACAGCTCATCATCAATCAAAACCGCCCGGAAAGCAAGATGGCGGCG 322
 QY 38 LysProGluGlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeu----- 52
 Db 323 TGCTCAACCTCAGCCCAAGCG---CCAGAGCAGAGGAGCGCTGCATCCCGAAGGGA 379
 QY 53 ---ArgAspValSerGluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyroCys 71
 Db 380 GCCCATGATGTCGGGAAGAGTGTGACGCGCACTGGAAGACATCATCAAGCAGTAT--- 436
 QY 72 ValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluPro 91

437 -----GGGTCTCTCCAGTCCCGAGGGAAG 463

92 GluAspAlaGluLysSerArgThrTyrValAlaArgAsnGlyGluProGluProThrPro 111
||| : : : : :
464 GAGAGCACCTCTGAAACTAAGGAGCAGCCCCCAACACAGAGGCACACAGCAATAGGAT 523
||| : : : : :
112 ValValTyrGlyGluLysGluProSerLysGlyAspProAsnThrGluGluIleArgGln 131
||| : : : : :
524 GTTGACTATGAAGA-----ACCACCTGAAGAGATA----- 553
||| : : : : :
132 SerAspGluValGlyAspArgAspHisArgArgProGln----- 144
||| : : : : :
554 -----GACAGAGAACCCACTGCTCTCAAGAGCCGCGCAGCCAAAGAG 598
||| : : : : :
145 -----GluLysLysLysAlaLysGlyLeuGlyLysGlu 155
||| : : : : :
599 CCTGTGAGCAATAAAGAGCAAAAGCTGGAAAAGAAATCTTAAAGAGATTAGTAAAGAA 658
||| : : : : :
156 IleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAla 175
||| : : : : :
659 GCTAACTTGCTCATGCAAAACCTGAAACAGCTGCAAGCACCTGAGNAAGCTTGATTT 718
||| : : : : :
176 LeuCysLysIleTyrAlaGluLeuLeuGluGluHisArgAsnSerGlnLysMetLys 195
||| : : : : :
719 TTATTCAAGAAGTATACTGAATTGCTGGATGAACATGCCACCGAGCAAAAGAACTCAAG 778
||| : : : : :
196 LeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHis 215
||| : : : : :
779 CTCTTACTGAACAGCAGGGCGAGCCAGAGAGAGAGGACCAGTTGCGAGTGAGCAC 838
||| : : : : :
216 SerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHis 235
||| : : : : :
839 AACAGACCGTCTCTGTCGAAGCAAACTCGAGAGTCTGTGTCGGAGCTACAGAGCCAC 898
||| : : : : :
236 AsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGluGluGluGluLysArgLys 255
||| : : : : :
899 AACAGAAGCGTGAAGGAGGAAACCTCCACCGGCGCGAGGAAGAGAGAGAGGAAA 958
||| : : : : :
256 GluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHis 275
||| : : : : :
959 GAGATCACAGTCACTTCTCAAACTACCTGACAGATATCCAGACTCAGATCGAACAACAG 1018
||| : : : : :
276 AsnGluArgAsnSerLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLys 295
||| : : : : :
1019 AGTGAGCGCAATATGAAGCTCTGCAGAGAGAACACAGAGCTTCGAGAAAACCTGAAAGC 1078
||| : : : : :
296 LeuIleGluGlnTyrGluLeuArgGluGluHisIleAspLysValPheLysHisLysAsp 315
||| : : : : :
1079 ATCATTGACCAAGTACGAGCTCAGAGAGAGCACTCGACAAATATTTCAAAACATCGAGAA 1138
||| : : : : :
316 LeuGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnLeuMetLeuLysGluAla 335
||| : : : : :
1139 CTGCAGCAGAAGCTGGCGTATGCAAACTTGAAGAGGCTCAAGAACTGATGAGGAGGCA 1198
||| : : : : :
336 GluGluArgHisGlnArgGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArg 355
||| : : : : :
1199 GAGAGCGGCACAGACGGGAGAGGAATATTGCTGAACAGCGCGCAGAGTGGAACTT 1258
||| : : : : :
356 MetCysGluLeuMetLysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrThr 375
||| : : : : :
1259 CAGGCCAAAGTGTGAAGGAGCAGAGCGTGTGCGAGGCTCAGGCTCAGCTCACTCTACTCA 1318
||| : : : : :
376 GluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPhe 395
||| : : : : :
1319 GGAAGGTTTGAGGAGTTCGAGATPACACTGACAAAAGCAATGAAGTGTTCGCACCTTC 1378
||| : : : : :
396 LysGlnGluMetGluLysMetThrLysLysIleLysLysLeuGluLysGluThrThrMet 415
||| : : : : :
1379 AAGCAGGAATGGATAGACAACTAAGAAATGAAGAGCTGGAGAGGACACCGCCACG 1438
||| : : : : :
416 TyrArgSerArgTrpGluSerSerAsnLysAlaLeuLeuMetAlaGluGluLysThr 435
||| : : : : :
1439 TGGAAAGCCCATTTGAATTTGAAACAAAGCCCTGTTGGAATGATCGAAGAGAAAGCC 1498
||| : : : : :

TITLE
JOURNAL

Direct Submission
Submitted (16-Apr-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
Location/Qualifiers
1. .2480
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:D830027D20"
/db_xref="MGI:2423165"
/db_xref="taxon:10090"
/clone="D830027D20"
/tissue type="heart"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="16 days neonate"
1. .2480
/note="MUSCLE-DERIVED PROTEIN MDP77 VARIANT 1 (SPTR|AAL33909, evidence: FASTV, 99.1%ID, 100%length, match=2058)"

FEATURES

source

misc_feature

ORIGIN

Alignment Scores:

Pred. No.:	3.35e-68	Length:	2480
Score:	1193.00	Matches:	265
Percent Similarity:	63.15%	Conservative:	76
Best Local Similarity:	49.07%	Mismatches:	139
Query Match:	44.15%	Indels:	60
DB:	11	Gaps:	8

US-10-023-523-8 (1-530) x AK085904 (1-2480)

Qy 18 ProSerGlnAlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArg 37
Db 156 CCCCAGGGGACAGCTCATCTCAATCAAAACGCGCGGAAAGCAAGATGGCGAGCGG 215
Qy 38 LysProGluGlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeu----- 52
Db 216 TGCTCAACCTCAGGCGCAAGCG---CCAGACAGAGGAGGAGCGCTGCATCCCGAGAGGGA 272
Qy 53 ---ArgAspValSerGluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrcys 71
Db 273 GCCCATGATGCGGAGAGAGTTGACGAGGCACTGGAGACATCATPAGCACCTAT--- 329
Qy 72 ValAspAsnAsnGlnGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluPro 91
Db 330 -----GGTCTGCTGCGCAGTCCCGGAGGGAAG 356
Qy 92 GluAspAlaGluLysSerArgThrTyrrValAlaArgAsnGlyGluProGluProThrPro 111
Db 357 GAGACACCTCTGAACCTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 416
Qy 112 ValValTyGlyGluLysGluProSerLysGlyAspProAsnThrGluGluLeuArgGln 131
Db 417 GTTGACTATGAAGAA-----ACCATGAAGAGATA----- 446
Qy 132 SerAspGluValGlyAspArgAspHisArgArgProGln----- 144
Db 447 -----GACAG 491
Qy 145 -----GluLysLysLysAlaLysGlyLeuGlyLysGlu 155

Db 492 CCTGTCAGCAATAAGAGCAAGCAAGCTGGAAGAAAGAAATCTTAAAGGATTAGTAAGAA 551
Qy 156 IleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAla 175
Db 552 GCTAACTTGTCTCATGCAAAACCTGAAACAGCTGCAAGCACCTGAAAGAAAGCTGATTT 611
Qy 176 LeuCysLysLysTyrrAlaGluLeuLeuGluGluHisArgAsnSerGlnLys-GlnMetly 195
Db 612 TTATTCAAGAGAGTATCTGAATGTGTGATGAACATCGCACCGAGCAAAAGAGCTGAA 671
Qy 195 sleuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGlu 215
Db 672 GCTCTCTACTGAACACAGCAGCGCAGACCCAGAGAGAGAGACCACTTGCAGAGTGAG 731
Qy 215 sSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGluArg 235
Db 732 CAACAGAGCGCTCTTGTCTGAAGCAAACTCGAGAGTCTGTGCGGAGCTACAGAGGCA 791
Qy 235 sAsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGluGluGluGluLysArg 255
Db 792 CAACAAGCGCTGAGAGAGAAACCTCCACGGCGCGGAGAGAGAGAGAGAGAGAGAG 851
Qy 255 sGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGln 275
Db 852 AGAGATCAACAAGTCACTTCAAACTACCTCGACAGATATCCAGACTTCAGATCGAACAA 911
Qy 275 sAsnGluArgAsnSerLysLeuArgGlnGluLeuMetGlu-LeuAlaGluArgLeuLys 295
Db 912 GAGTGAGCGCATATGAGCTCTGCCAGGAGAACACAGAGCCCTTCAGAAAATGAGAA 971
Qy 295 ysLeuIleGluGlnTyrrGluLeuArgGluGluHisIleAspLysValPheLysHis 315
Db 972 GCATCATTCACCACTACGAGCTCAGAGAGAGAGATCTCGACAAAATATTCAACATCGAG 1031
Qy 315 sPleuGlnGlnLeuValAlaAspAlaLysLeuGlnGlnAlaGlnGlnMetLeuLysGlu 335
Db 1032 AACTGCGCAGAGAGCTGTGTGATGCGAACTTGAAGAGGCTCAGAAGCTGTATGCGAGG 1091
Qy 335 laGluGluArgHisGlnArgGluLysAspPheLeuLysGluAlaValGluSerGlnA 355
Db 1092 CAGAGAGCGGCACAGACGCGGAGAGAGATATTTCTGAACAGCGCGCAGAGTGGAAAC 1151
Qy 355 rgMetCysGluLeuMetLysGlnGlnThrHisLeuLysGlnGlnLeuAlaLeuTyrr 375
Db 1152 TTCAGGCCAAAGTGTGAAGAGCAGGACGCGTCTGCAGGCTCAGCTCTCTCTACT 1211
Qy 375 hrCluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluValPheThrThr 395
Db 1212 CAGGAAGGTTTGGAGAGTTCCAGAGTACACTGCAAAAAGCAATGAAGTGTTCCTACT 1271
Qy 395 helysGlnGluMetGluLysMetThrLysLysIleLysLysLeuGluLysGluThrThr 415
Db 1272 TCAAGCAGGAATCGATTAAGACAACTAAGAAATGAAGAGCTGGAGAGAGACCGCCA 1331
Qy 415 etTyrrArgSerArgTrpGluSerSerAsnLysAlaLeuLeuGluMetAlaGluLys 435
Db 1332 CGTGGAAAGCCCGATTTGAGAAATTGTAACAAAGCCCTGTGTGACATGATCGAAGAGAAG 1391
Qy 435 hrValArgAspLysGluLeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeu 455
Db 1392 CCTCGCGACTAAAGAAATACGAGTGTTCGTGATGAAATCCAAAGGCTGGAGAACCTGT 1451
Qy 455 yaArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAla 475
Db 1452 GTGCGTCTTACAGAGGAGAGAAAGCACTCTAAGAAATCAGAGAGAGCAAAATGT 1511
Qy 475 lyGlyGlnGlnSerLeuThrAspSerGlyProGluArgProGluGlyProGlyAla 495
Db 1512 CCGAAAG 1568
Qy 495 InAlaProSerSerProArgValThrGluAlaProCysTyrrProGlyAlaProSerThr 515

```

Db      1569 CTGAAATGAAGAGGTGGAGCGAGAGGAGCCACAGATTTTCAAAAAGCTGTGGAACCC 1628
Qy      515 luAlSerGlyGlnThr-----GlyProGlnGluProThr 526
Db      1629 TGCCACACAGCTTTTACCATCTTCATCCAGAGTTCACCCCGACCAACCAACG 1684

RESULT 18
CA327097
LOCUS   CA327097 750 bp mRNA linear EST 09-JUL-2003
DEFINITION UI-W-FY0-cy-g-01-0-UI-r1 NIH_BMAP_FY0 Mus musculus cDNA clone
IMAGE:6826106 5', mRNA sequence.
ACCESSION CA327097
VERSION CA327097
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 750)
JOURNAL NIH-MGC http://mgi.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
Location/Qualifiers
1..750
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6826106"
/tissue_type="whole brain"
/dev_stages="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_FY0"
/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGCGAGACAG. This library was created for the University
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."
ORIGIN
Alignment Scores:
Pred. No.: 9.19e-67 Length: 750
Score: 1161.50 Matches: 231
Percent Similarity: 93.98% Conservative: 3
Best Local Similarity: 92.77% Mismatches: 14
Query Match: 42.99% Indels: 1
Gaps: 1
DB: 14

US-10-023-523-8 (1-530) x CA327097 (1-750)
Qy      45 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 64

```

```

Db      2 AAAGCACTCAGCTGGGGCCCTCTGTGACGTCTCTGAGAGCTGAGCCGCGAGTTGGAA 61
Qy      65 AspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyVala 84
Db      62 GACATCTGAGTACATACCTGTGTGGACAACAATCAGGAGGCCCGCTGAGGAGGAGCA 121
Qy      85 GlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrValAlaAlaArgAsn 104
Db      122 CAGGGTGGAGCCCATGAGCCGGAAGACACGAGAGAAGTCCCGAACTATGATGAGCCAGGAAT 181
Qy      105 GlyGluProGluPro---ThrProValValTyrGlyGluLysGluProSerLysGlyAsp 123
Db      182 GGGGAGCTGACACGAGCATTCAGTCTGTCACGCGAGAGGAGAGACCTCTAAGGGAGG 241
Qy      124 ProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAspHisArgArgPro 143
Db      242 CCTGGAACAGAGAGATCCGAGCGAGTGTATGAAGTTGGAGACCGAGACCATCGAGGCCA 301
Qy      144 GlnGlyLysLysAlaLysGlyLeuGlyGlyGluLeuThrLeuLeuMetGlnThrLeu 163
Db      302 CAGGAGAAGAAAGCAAGGGTCTAGGGAAGAGATCACTCTGCTGATGACAGACTG 361
Qy      164 AsnThrLeuSerThrProGluGluLysLeuAlaAlaLeuCysLysLysTyrAlaGluLeu 183
Db      362 AACAGCTGAGTACCCAGAGAGAGAGCTGGCTGCATGTGCAAGAGATATGCTGAGCTG 421
Qy      184 LeuGluGlnHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGln 203
Db      422 CTGGAAGAGCATCGAACTCGCAAGACGATGAAGCTCCTGCAGAGAAGAGAGAGCCAG 481
Qy      204 LeuValGlnGlnLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSer 223
Db      482 CTGTCAGAGAGAGGACCATCTCGAGGGGAACACAGCAAGAGCTGTCTGGCCCCGAGG 541
Qy      224 LysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGly 243
Db      542 AAGCTTGAGAGTCTGTCCGGAGCTGCAACGCGCACACCCGTCCTTGAAGGAAGAGGC 601
Qy      244 ValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnVal 263
Db      602 GTGACGAGCCCGTGGAGGAGGAGAGCGGCANAGAGTGTCTTCCACTTCCAGGTG 661
Qy      264 ThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGlnLysArgAsnSerLysLeuArg 283
Db      662 ACATGAATGACATTCAGCTGCAGATGGAACAGCATACGAGCGANACTCCAGTCCGCG 721
Qy      284 GlnGluAsnMetGluLeuAlaGluArg 292
Db      722 CAGGAGAAATATGAGCTAGCCGAGAGG 748

RESULT 19
LOCUS   BQ714219 942 bp mRNA linear EST 16-JUL-2002
DEFINITION AGENCOURT_8291711 NIH_MGC_129 Mus musculus cDNA clone IMAGE:6308909
5', mRNA sequence.
ACCESSION BQ714219
VERSION BQ714219
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 942)
JOURNAL NIH-MGC http://mgi.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Susan L. Sullivan, PhD.
cDNA Library Preparation: ResGen, Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation

```


Alignment Scores:

Pred. No.: 2,75e-66 Length: 727
 Score: 1154.00 Matches: 232
 Percent Similarity: 97.91% Conservative: 2
 Best Local Similarity: 97.07% Mismatches: 5
 Query Match: 42.71% Indels: 1
 DB: 13 Gaps: 0

US-10-023-523-8 (1-530) x BQ770181 (1-727)

QY 111 ProValValTyrGlyGluLysGluProSerLysGlyAspProAsnThrGluGluLeuArg 130
 Db 7 CCAGTCGTCAACGGGAGAGAGACCTCTTAGGGAGAGCCCTGGAAACAGAGAGATCCGA 66
 QY 131 GlnSerAspGluValGlyAspArgAspHisArgArgProGlnGluLysLysLysAlaLys 150
 Db 67 CGAGTCGTATGAAGTTGGAGACGAGACCATCGAGGCCACAGGAGAGAAAGCCCAAG 126
 QY 151 GlyLeuGlyLysGluLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGlu 170
 Db 127 GGTCTAGGAGAGAGATCACTCTGCTGATGCAGACACTGAACAGCTGAGTACCCAGAG 186
 QY 171 GluLysLeuAlaAlaLeuCysLysLysTyrAlaGluLeuLeuGluGluHisArgAsnSer 190
 Db 187 GAGAGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 246
 QY 191 GlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHis 210
 Db 247 CAGAAGCAGATGAAGCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 306
 QY 211 LeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArg 230
 Db 307 CTGGAGGGGAGACAG 366
 QY 231 GluLeuGlnArgHisAsnArgSerLysGluGlyValGlnArgAlaArgGluGlu 250
 Db 367 GAGCTGCAACGGCACAACCGTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 426
 QY 251 GluGlnLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspLeuGlnLeu 270
 Db 427 GAGGAGAGAGCGCAAG 486
 QY 271 GlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGluLeuAla 290
 Db 487 CAGATGCAACAGCATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 546
 QY 291 GluArgLysLysLysLeuLeuGluGlnTyrGluLeuArgGluGluHisLysLysVal 310
 Db 547 GAGAGGCTCAAGAGTTGATCGAGCAATACGAGCTTCGTGAGGAGCATATCGACAAAGTC 606
 QY 311 PheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGlu 330
 Db 607 TTTCAACATAGGACCTGCGAGCAGCAGCTAGTGACGCCAAGCTCCAGAGGCCAGGAG 666
 QY 331 MetLeuLysGluAlaGluGluArgHisGlnArgGluLysAspPheLeuLysGlu 349
 Db 667 ATGCTGAAGGAGGAG 722

RESULT 21

CF728006
 LOCUS 716 bp mRNA linear EST 09-OCT-2003
 DEFINITION UI-M-HB0-ck1-n-24-0-UI.r1 NIH BMAP_HB0 Mus musculus cDNA clone
 IMAGE:30550511 5', mRNA sequence.

CF728006

CF728006.1 GI:37602174

EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 716)

NIH-MGC http://mgi.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE

JOURNAL COMMENT

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapsb@mail.nih.gov

Tissue Procurement: Dr. James Lin University of Iowa

CDNA Library preparation: Dr. M. Bento Soares, University of Iowa

CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

CDNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Distribution information can be found at

http://genome.uioawa.edu/distribution/mousef.html

This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)

Seq primer: pyX-5.

Location/Qualifiers

1. 716

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAG:30550511"

/tissue_type="whole eye"

/dev_stage="embryo 12.5,13.5,14.5 dpc"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NIH_BMAP_HB0"

/notes="Organ: Eye; Vector: pyX- Asc; Site 1: EcoR I;

Site 2: Not I; The library was constructed according

Bonaldi, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT

size selected according to mRNA size fraction, ligated

with EcoR I adaptor, digested with NotI and then cloned

directionally into pyX-Asc vector. The library tag

sequence located between the Not I site and the polyA tail

is TTTATGAGT. This library was created for the University

Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the

Developing Mouse Nervous System', supported by National

Institute of Mental Health (NIMH)."

ORIGIN

Alignment Scores:

Pred. No.: 2,92e-66 Length: 716
 Score: 1153.50 Matches: 232
 Percent Similarity: 97.51% Conservative: 3
 Best Local Similarity: 96.27% Mismatches: 3
 Query Match: 42.69% Indels: 3
 DB: 14 Gaps: 1

US-10-023-523-8 (1-530) x CF728006 (1-716)

QY 250 GluGluGlyLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspLeuGln 269
 Db 1 GAGGAGGAGAGCGCAAG 60
 QY 270 LeuGlnMetGluGlnHisAsnGluArgSerLysLeuArgGlnGluAsnMetGluLeu 289
 Db 61 CTGCGAGATGGAACAGCATACAGGCGAACTCCAGCTGCCAGGAGAAATATGAGCTA 120
 QY 290 AlaGluArgLysLysLysLeuLeuGlnTyrGluLeuArgGluGluHisLysLys 309
 Db 121 GCCGAGAGGCTCAAGAGTTGATTCAGCAATACAGAGTTCGTGAGGAGCATATCCACAA 180
 QY 310 ValPheLysHisLysAspLeuGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGln 329
 Db 181 GTCTTCAACATAGGAGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
 QY 330 GluMetLeuLysGluAlaGluGluArgHisGlnArgGluLysAspPheLeuLysGlu 349
 Db 241 GAGATGCTGAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
 QY 350 AlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLeuLysGln 369
 Db 301 GCGGTGGAGTCCAGAGGATGTCAGCTGATGATGAGAGAGAGAGAGAGAGAGAGAGAG 360

QY 370 GlnLeuAlaLeuThrThrGluLysPheGluGluPheGlnAsnThrLeuSerIysSerSer 389
 DB 361 CAGTCGCCCTGTACACGGAGAGTTTGGAGAGTTCCAGAACACACTTTCGAAAGCAGT 420
 QY 390 GluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysIleLysLysLeu 409
 DB 421 GAAGTGTTCACACAGTTCACACAGGAGATGGAAGAGATGACAAAGATCAAGAGCTG 480
 QY 410 GluLysGluThrThrMetThrArgSerArgTTPGluSerSerAsnLysAlaLeuLeuGlu 429
 DB 481 GAGAGAAACACCACTATGATCCCGTGGAAAGCGCAAGGCTCTGCTGGAG 540
 QY 430 MetAlaGluLysThrValArgAspLysGluLeuGluGluValLysIleGln 449
 DB 541 ATGGCTGACGAGAAACCGTCCGGGACAAAGAGCTGGAGGGCTGCAGGTGAAATCCAG 600
 QY 450 ArgLeuLysLeuLysCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgVal 469
 DB 601 CGCTGGAGAGCTGTGCGGAGCACTGCAGACCGGCGCATGACCTCAACAGAGGGTA 660
 QY 470 GlnAspLeuSerAlaGlyGlnGlySerLeuThrAspSerGlyProGluArgPro 489
 DB 661 CAGGACCTGACTCAGGGGCGC-----ATCAGTACATTGGCTCTGAGCGGAGGCA 711
 QY 490 Glu 490
 DB 712 GAG 714

RESULT 22
 BF783468
 LOCUS 602111239F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:4239513
 DEFINITION 5' mRNA sequence.
 ACCESSION BF783468
 VERSION EST.
 KEYWORDS BF783468.1 GI:12088504
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 819)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs@mail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLN9852 row: 1 column: 10
 High quality sequence stop: 758.
 Location/Qualifiers
 1..819
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4239513"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NCI CGAP Kid14"
 /note="Organ: kidney; Vector: pCMV-SPORT6; Site 1: Not I; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.75 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP library."|

FEATURES

source
 1..819
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4239513"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NCI CGAP Kid14"
 /note="Organ: kidney; Vector: pCMV-SPORT6; Site 1: Not I; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.75 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP library."|

ORIGIN

Alignment Scores: 2.46e-65 Length: 819
 Pred. No.: 1140.50 Matches: 239
 Score:

Percent Similarity: 91.67% Conservative: 3
 Best Local Similarity: 90.53% Mismatches: 18
 Query Match: 42.21% Indels: 5
 DB: 10 Gaps: 1
 US-10-023-523-8 (1-530) x BF783468 (1-819)

QY 59 LeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGly 78
 DB 2 CTGAGCCGCGCAGTTGGAAGACATCTCTGATACATCTGTGTGGACACATCATCGGAGGC 61
 QY 79 ProGlyGluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArg 98
 DB 62 CCGGCTGAGGAGGAGACACAGGGTGGCCCTGAGCCGGAAGACACGCGAAGTCCCGA 121
 QY 99 ThrTyrValAlaArgAsnGlyGluProGluPro---ThrProValValTyrGlyGluLys 117
 DB 122 ACTATGAGCCAGGATGGGAGGCTGACACAGGCATTCAGCTGTCACAGCGGAGAG 181
 QY 118 GluProSerLysGlyAspProAsnThrGluGluLeuArgGlnSerAspGluValGlyAsp 137
 DB 182 GAGACCTCTAAGGAGAGCCTGGAACAGAGAGATCCGACGACGATGATGAGGTGGAGAC 241
 QY 138 ArgAspHisArgArgProGlnGluLysLysAlaLysGlyLeuGlyLysGluLeuThr 157
 DB 242 CGAGACCATCGGAGGCCACAGGAGAGAGAAAGCCAGGGTCTAGGGAGGAGATCACT 301
 QY 158 LeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCys 177
 DB 302 CTGCTGATGACACACTCAACACGCTGATACCCAGAGAGAGCTGGCTGCCTGCTGC 361
 QY 178 LysLysTyrAlaGluLeuGluHisArgAsnSerGlnLysGlnMetLysLeuLeu 197
 DB 362 AAGAAGTATCTGAGCTCTGAGAGAGATCGAACTCCGCAAGACAGATGAAGCTCTG 421
 QY 198 GlnLysLysGlnSerGlnLeuValGlnGlnLysAspHisLeuArgGlyGluHisSerLys 217
 DB 422 CAGAAGAGCAGAGAGCAGCTGTCGAGGAGAGAGACCATCTCGAGGGGAAACAGCAAG 481
 QY 218 AlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArg 237
 DB 482 GTGTCTGCGCCGAGCAAGCTTGAGAGTCTGTGCGGAGAGCTGCAACGCGCAACCCG 541
 QY 238 SerLeuLysGluGluGlyValGlnArgAlaArgGluGluLysArgLysGluVal 257
 DB 542 TCCCTGAAGAGAGAGCGCTGCGAGCCGAGAGAGGAGAGAGCGCAAGAGATG 601
 QY 258 ThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGlu 277
 DB 602 ATTCACACTTCCAGTGACACTGAATGACATTCAGCTGAGATGGAGACAGCATTAACGAG 661
 QY 278 ArgAsn-SerLysLeuArgGlnGluAsnMetGluLeuAlaGlu-ArgLeu-LysLysLeu 286
 DB 662 CGAAACCTCCAAAGCTGGGCGGAGAGATATGGAGCTAGCCGAGAGGCTCAAAAGATTG 721
 QY 297 IleGluGlnTyrGluLeuArgGluLysHisIleAspLysValPheLysHisLysAspLeu 316
 DB 722 ATCGAGCAATCGAGAGCTTTCGAGAGAGCTTTTCGACAGAGTCTT-AAACATTAGGACTGC 780
 QY 317 GlnGln 318
 DB 781 GAGCAG 786

RESULT 23

AK031023

LOCUS

DEFINITION

AK031023.1

GI:26082107

KEYWORDS

HTC; CAP trapper.

SOURCE

Mus musculus (house mouse)

AK031023 4452 bp mRNA linear HTC 18-SEP-2003
 Mus musculus adult male thymus cDNA, RIKEN full-length enriched library, clone:5830498L23 product:hypothetical protein, full insert sequence.

```

ORGANISM      Mus musculus
REFERENCE     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS       Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE         1. Carninci, P. and Hayashizaki, Y.
JOURNAL       High-efficiency full-length cDNA cloning
MEDLINE       Meth. Enzymol. 303, 19-44 (1999)
PUBMED       99279253
PUBMED       10349636
REFERENCE     2.
AUTHORS       Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
              Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE         Normalization and subtraction of cap-trapper-selected cDNAs to
              prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL       Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE       20499374
PUBMED       11042159
REFERENCE     3.
AUTHORS       Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
              Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
              Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
              Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
              Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
              Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
              Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE         RIKEN integrated sequence analysis (RISA) system--384-format
              sequencing pipeline with 384 multicapillary sequencer
JOURNAL       Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE       20530913
PUBMED       11076861
REFERENCE     4.
AUTHORS       The RIKEN Genome Exploration Research Group Phase II Team and the
              FANTOM Consortium.
TITLE         Functional annotation of a full-length mouse cDNA collection
JOURNAL       Nature 409, 685-690 (2001)
REFERENCE     5.
AUTHORS       The FANTOM Consortium and the RIKEN Genome Exploration Research
              Group Phase I & II Team.
TITLE         Analysis of the mouse transcriptome based on functional annotation
              of 60,770 full-length cDNAs
JOURNAL       Nature 420, 563-573 (2002)
REFERENCE     6.
AUTHORS       Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
              Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
              Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
              Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
              Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
              Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
              Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
              Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
              Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
              Segabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
              Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
              Muramatsu, M. and Hayashizaki, Y.
TITLE         Direct Submission
JOURNAL       Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
              Physical and Chemical Research (RIKEN), Laboratory for Genome
              Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
              RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
              Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
              URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
              Fax: 81-45-503-9216)
COMMENT       cDNA library was prepared and sequenced in Mouse Genome
              Encyclopedia Project of Genome Exploration Research Group in Riken
              Genomic Sciences Center and Genome Science Laboratory in RIKEN.
              Division of Experimental Animal Research in Riken contributed to
              prepare mouse tissues.
              Please visit our web site for further details.
              URL: http://genome.gsc.riken.go.jp/
              URL: http://fantom.gsc.riken.go.jp/
              Location/Qualifiers
              1. 4452
                 /organism="Mus musculus"
FEATURES      source

```

```

/mol_type="mRNA"
/db_xref="PANTOM_DB:5830498L23"
/db_xref="WGI:2394192"
/db_xref="taxon:10090"
/clone="5830498L23"
/sex="male"
/tissue_type="thymus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
1. 4452
misc_feature
ORIGIN
Alignment Scores:
Pred. No.:      3,556-64      Length:      4452
Score:          1136.50      Matches:    260
Percent Similarity: 43.97%      Conservative: 68
Best Local Similarity: 34.85%      Mismatches: 67
Query Match:      42.06%      Indels:     351
DB:              11          Gaps:       5
US-10-023-523-8 (1-530) x AK031023 (1-4452)
QY      19 SerGlnAlaAlaProAlaValGluAlaGluGlyPro-----GlySerSerGlnAlaPro 36
Db      958 TCTCAGCAAGTAATGATATTCTTCACATCAAGACCCAGTGTGGTGGCACAACTAAGAA 1017
QY      37 ArgLysProGluGlyAlaGlnAlaAargThrAlaGlnSerGlyAlaLeuArgAspValSer 56
Db      1018 CATTCACCTGGAAGGGGATGAAGGC-----AGT 1044
QY      57 GluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysVal---AspAsnAsn 75
Db      1045 GACTTTATTACAAAGAACAGAAATTTGGTGAGCTCAGTATTCTGTACACAGAGAAAGA 1104
QY      76 GlnGlyGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGlu 95
Db      1105 GAAGAAATTCCTGGACGGGAGCTCGAACAGTCTCTCTGATGGCCAGCAAGATTACAG 1164
QY      96 LysSerArgThrTyrValAlaAargAsnGlyGluProGluProThrProValValTyrGly 115
Db      1165 TGCACGACG----- 1173
QY      116 GluLysGluProSerLysGlyAspProAsnThrGluGluIleArgGlnSerAspGluVal 135
Db      1173 ----- 1173
QY      136 GlyAspArgAspHisArgProGlnGluLysLysLysAlaLysGlyLeuGlyLysGlu 155
Db      1174 -----AACAAAGAGAGAGACCTTAGGAAAGAA 1200
QY      156 IleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaAla 175
Db      1201 GTTTATTACTATGATCAAGCGGTAAACACCCCTTTCAACCCACAGAGAGAGCTGCAGCT 1260
QY      176 LeuCysLysLysTyrAlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLys 195
Db      1261 CTCCTGAAGAAATATGCTGATCTCTCGAAGAAAGACAGGAATGTTTCAGAAACAAATGAAG 1320
QY      196 LeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHis 215
Db      1321 ATTCGCAAGAAAGAACCAAGCCAGATTGTGAAGAGAAAGATTTCACCTTCAGAGTGAACAC 1380
QY      216 SerLysAlaValLeuAlaAargSerLysLeuGluSerLeuLysArgGluLeuGlnArgHis 235
Db      1381 ACCAAGGCCATCTTGGCAAGAGCAAACTGGAATCTCTTTGCAGGGAACCTTCAGCGTCAAT 1440
QY      236 AsnArgSerLeuLysGluGlyValGlnArgAlaAargGluGluGluGluLysArgLys 255
Db      1441 AATAAGACCTTAAGAGGAGAGAAATATGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500

```

```

QY 256 GluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHis 275
DB 1501 GAAGCAACAGCAGCATTTCCGATAACTCTAAATGAATCCAAAGCTCAGTTGGACAACAT 1560
QY 276 AsnGluArgAsnSerLysLeuArgGlnGlnGlnMetGluLeuAlaGluArgLeuLysLys 295
DB 1561 GACATCCACATGCCAACTCCGACAGGAGAACATGTAACCTGGAGAGAGAGTTGGAAGAG 1620
QY 296 LeuLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 315
DB 1621 CTTATTGACAGTATGCACTAAGGCAAGAGCATATTGATTAAGTATTCAAAACACAGGAA 1680
QY 316 LeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 335
DB 1681 TTGCACCAACAGCTGTGGATGCCAACTTCAGCAACAACACAGCTGATTAAGAAGCT 1740
QY 336 GluGluArgHisGlnArgGluLysAspPhe 345
DB 1741 GATGAACCAACATCAGAGAGAGAGAGAGTTTGTAAAGTTCTGTTCTTTAAAAAAGGAAA 1800
QY 345 345
DB 1801 AGTATTGCCAGTATTTTAACTAATTCATGTCATCTCTGAGGTATGATGAAAAGCCCT 1860
QY 345 345
DB 1861 GGAAGGCTTCTGGCTACTATTCTTATTCAGAGATCCACATGCCGTGTTGAGTGCTAGGA 1920
QY 345 345
DB 1921 TTAAGAGGATGCATTACCATTCCTCCCTGCTTTAGTTTGTGAACAAGTTAAACACTA 1980
QY 345 345
DB 1981 GGGAAATGATTTTCATGGTGAATTTTAATTTATTTAGTATCATTTCTCCCTTTATTT 2040
QY 345 345
DB 2041 TTGTTGATCAGGCTCATATTGATGCTCTGGCTGGGGATGACCTTGAACACTTGAA 2100
QY 345 345
DB 2101 GCACATCCCAATTTAGGAAGTCTCGGGATTAAACCTCGGGCTTTCTACTTACATATA 2160
QY 345 345
DB 2161 AAGCAAGCAGATCTCCAGTCTCACACCATTTTGTGTTCTCTAGACTTTTACTACTTT 2220
QY 345 345
DB 2221 CCTCCATAGCCTTTGCGCCCATCTCTCTGTTCTCTTAGGTAGTAGAGAGGTTTCTCAG 2280
QY 345 345
DB 2281 TCATGTAGGAGTACTTAGCCATGTTTCTGCTAATATCAGACATGTGATTTATTTTACC 2340
QY 345 345
DB 2341 TGAACCAACCAACACACATCTTTAAACCTCAACTTTGTAATTTGTTTAGGCTCCAC 2400
QY 345 345
DB 2401 TATACCTTTCTACTTAAGCAGTAAATAATATAAGTATTCAAAACATTTCTACTTCTTAA 2460
QY 345 345
DB 2461 TGATGGAAGGTATTGTAAATGGTAACTGGTGGAAATTAAGGTGTTCATTTTGGAGGTA 2520
QY 345 345
DB 2521 GTCATTTCTGTAATATGATGAGTGAAGATGTGGAGTAACATGATGTTGATCTGACATAT 2580
QY 346 346

```

```

DB 2581 TATGTTCTTATTTGTAATTCCTATTAGTACTTTTCCAAAGATGTTTCTATTATTTAGTTA 2640
QY 347 LeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlnThrHis 366
DB 2641 TTAAGAAGCAACAGATCCAGGCACAAATATGACAAATGAACACAGACAGAGTACAA 2700
QY 367 LeuLysGlnGlnLeuAlaLeuThrThrGluLysPheGluGluPheGlnAsnThrLeuSer 386
DB 2701 CTAACACAGCAGCTTCTCTTTATATGATAAATTTGAAGAATTTCCAGACTACTATGCA 2760
QY 387 LysSerSerGluValPheThrThrPheLysGlnGlnMetGluLysMetThrLysLysIle 406
DB 2761 AAAGCAATGAACCTTTTACACCTTCAGGCAGGAATGGAAGATGACAAAGAAAT 2820
QY 407 LysLysLeuGluLysGluThrThrMetTyrArgSerArgTTPGluSerSerAsnLysAla 426
DB 2821 AAAAATCGAAAAAGAAACAATAATATGGCGTACCAATGGAAAAACAATAATAAGCA 2880
QY 427 LeuLeuGluMetAlaGluLysThrValArgAspLysGluLeuGluGluGlnVal 446
DB 2881 CTTCTGCAGATGCCGAAGAGAAAACTGTCCGTGATAAAGAGTACAAGGCTTTTCAAATA 2940
QY 447 LysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsn 466
DB 2941 AAATCGAACGTTAGAGAGCTGTGCAGGCTCTTTCAGACAGAGAGAAATGAGCTCAAC 3000
QY 467 LysArgValGlnAspLeu 472
DB 3001 GAGAAGGTGGAAGTCTG 3018

```

RESULT 24

CA324134

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CA324134 778 bp mRNA linear EST 09-JUL-2003
 UI-M-F10-cco-f-03-0-UI-r1 NIH_EMAP_F10 Mus musculus cDNA clone
 IMAGE:6822244 5', mRNA sequence.

CA324134
 CA324134.1 GI:24542232

EST.
 Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

NIH-MGC <http://mgi.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

This clone was contributed by the Brain Molecular Anatomy Project
 (BNAP)

Seq primer: pYX-5.
 Location/Qualifiers

1. 778
 /organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:6822244"

/tissue_type="whole brain"

/dev_stage="embryo 13.5 14.5 16.5 17.5 dpc"

/lab_host="DH10B (TI phage resistant)"

/note="Organ: Brain; Vector: pYX-Asc; Site: 1; Ecor I;
 Site 2: Not 1; The library was constructed according
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCGAGACAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:

Pred. No.: 1.53e-64 Length: 778
Score: 1128.00 Matches: 232
Percent Similarity: 89.39% Conservativeness: 4
Best Local Similarity: 87.88% Mismatches: 17
Query Match: 41.75% Indels: 11
DB: 14 Gaps: 3

US-10-023-523-8 (1-530) x CA324134 (1-778)

QY 17 ArgProSerGlnAlaLeuProAlaValGluAlaGluGlyProGlySer----- 32
DB 3 CGACCAGACAGACGCTCT-----GGGGCGGAGCGGAGGTTCCACA 47
QY 33 SerGlnAlaPro-ArgLysProGluGlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLe 52
DB 48 AGCAGGCTCTCGGGAGAGACCGAGGGGCTCGAGCTAAAGCAGCTCAGCCTGGGGCCCT 107
QY 52 uArgAspValSerGluGluLeuSerArgGlnLeuAspIleLeuSerThrTyrCysVa 72
DB 108 CTGTGACGCTCTGAGGAGCTGAGCGGCGAGTTCGAGACATCTGAGTACTACTGTGT 167
QY 72 lAspAsnGlnGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluProG 92
DB 168 GGACAAACATCAGGAGGCGCGCTGAGGAGGAGCAGGAGCTGAGCCACTGAGCGCGGA 227
QY 92 uAspAlaGlySerArgThrTyrValAlaArgAsnGlyGluProGluPro---ThrPr 111
DB 228 AGACACGAGAGAGTCCCGAAGCTATGACAGGAGGAGTGGGAGGCTGACACGAGCATTC 287
QY 111 oValValTyrGlyGluLysGluProSerLysGlyAspProAsnThrGluGluIleArgG 131
DB 288 AGTGTCTAACGGCGAGAGGAGACCTCTAAGGAGAGGCTCGAACAGAGGAGATCCGAGC 347
QY 131 nSerAspGluValGlyAspArgAspHisArgArgProGlnGluLysLysLysAlaLysG 151
DB 348 GAGTGATGAAGTTGGAGACCGAGACCATCGGAGGCCAGGAGAGAGAGAAAGCCAAAGG 407
QY 151 yLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluG 171
DB 408 TCTAGGAGAGGAGATCATCTGCTGATGACAGACATGACAGCTGAGTACCCAGAGGA 467
QY 171 uLysLeuAlaAlaLeuCysLysLysTyrAlaGluLeuLeuGluLysArgAsnSerG 191
DB 468 GAACTGCTGCTGCTGCAAGAGATGCTGAGCTGCTGCAAGAGATCGCAACTCGCACTGCA 527
QY 191 nLysGlnMetLysLeuLeuGlnLysLysGlnSerClnLeuValGlnGluLysAspHisLe 211
DB 528 GAAGCAGATGAAGCTCTCGCAAGAGAGCAGGAGCAGCTCGTGAGGAGAGGACCATCT 587
QY 211 uArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgG 231
DB 598 GCGAGGGGACACAGAGGCTGCTCGCCCGAGAGAGCTTGAGAGCTCTGTGCGCGGA 647
QY 231 uLeuGlnArgHisAsnArgSerLeuLysGluGlyValGlnArgAlaArgGluGluG 251
DB 648 GCTGCAACGGCACAACCGCTCCCTGAGAGAGAGAGCGCTGAGCGAGCCCGGTGAGGAGGA 707
QY 251 uGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuG 271

Db 708 GGAGAGCCCAAGAGAGTGACCTTCACCTTCAGGTCGACATGATGACATTCACCTGCA 767
QY 271 nMetGluGln 274
Db 768 GATGGAACAG 777

RESULT 25

CB244426

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

CB244426

QY 95 GluLysSerArgThrTyrValAlaArgAsnGlyGluProGluPro---ThrProValVal 113
 Db 6 GAGAAAGTCCCGAACCTATGACGCCAGGAATGGGAGCCTGAACACAGGCAATCCAGTCGTC 65
 QY 114 TyrGlyGluLeuGluProSerLysGlyAspProAsnThrGluLeuLeuArgGlnSerAsp 133
 Db 66 AACGGCGAGAGAGACCTCTTAAGGAGAGAGCCTGGACAGAGAGATCCGAGCGAGTAT 125
 QY 134 GluValGlyAspArgAspHisArgArgProGlnGluLysLysLysAlaLysGlyLeuGly 153
 Db 126 GAAGTTGGAGACCGAGACCATCGAGGCCACAGGAGAGAGAAAGCAAGCGTCTAGGG 185
 QY 154 LysGluLeuThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeu 173
 Db 186 AAGGAGATCAGCTCTGCTGATGAGACATCTGAACACGCTGAGTACCCAGAGAGAGAGCTG 245
 QY 174 AlaAlaLeuCysLysLysTyrAlaGluLeuLeuGluGluHisArgAsnSerGlnLysGln 193
 Db 246 GTTGCACTGTGCAAGAGTATGCTGAGCTGCTGGAGAGCATCGGAATCTCGAGAGCGAG 305
 QY 194 MetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGly 213
 Db 306 ATGAGACTCTCTCAG 365
 QY 214 GluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGln 233
 Db 366 GAACACAGCAGAGCTGTCTGCTGCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 425
 QY 234 ArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGluGluGluLys 253
 Db 426 CGGCACACCGGCTCCTGAAG 485
 QY 254 ArgLysGluValThrSerHisPheGln-ValThrLeuAsnAspIleGlnLeuMetGln 273
 Db 486 CCAAGAAGAGTCACTTCACATTCACCGGTGACATCGAATGACATTCAGCTGCAGAGTGA 545
 QY 273 uGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLe 293
 Db 546 ACAGCATACAGAGCGAACTCAACGCTCGCCAGGAGAGATATGGAGCTAGCCGAGAGGCT 605
 QY 293 uLysLysLeuLeuGluGlnTyrGluLeu-ArgGluGluHisLeuAspLysValPheLysH 313
 Db 606 CAAGAAGTTGATCGAGCAATACGAGCTCCGCTGAGGAGCATATCGACAAAGCTCTCAAC 665
 QY 313 isLysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuL 333
 Db 666 ATAAGCACTGCAGAGCAGCTAGTGGAGCGCAGAGCTCCAGAGCGCCAGAGAGATGCTGA 725
 QY 333 YS 333
 Db 726 AG 727
 RESULT 26
 LOCUS BU701604
 DEFINITION UI-M-F10-Byp-e-23-0-UI-r1 NIH BMAP_F10 Mus musculus cDNA clone
 IMAGE:5721214 5', mRNA sequence.
 ACCESSION BU701604
 VERSION BU701604.1 GI:23625575
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 699)
 NIH-MGC <http://mgi.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 EMAIL: c9apbs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.

FEATURES
 source
 Location/Qualifiers
 1..699
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:5721214"
 /tissue_type="whole brain"
 /dev_stage="embryo 12.5dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH BMAP F10"
 /note="Organ: Brain; Vector: pYX-Asc; Site1: EcoR I;
 Site2: Not I; The library was constructed according
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with NotI and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is CAGCCACGAC. This library was created for the University
 Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."

ORIGIN

Alignment Scores:
 Pred. No.: 8,28e-64 Length: 699
 Score: 1116.00 Matches: 226
 Percent Similarity: 97.85% Conservative: 2
 Best Local Similarity: 97.00% Mismatches: 5
 Query Match: 41.30% Indels: 1
 DB: 13 Gaps: 0
 US-10-023-523-8 (1-530) x BU701604 (1-699)
 QY 115 GlyGluLysGluProSerLysGlyAspProAsnThrGluLeuLeuArgGlnSerAspGlu 134
 Db 1 GGGAGAGAGAGACCTCTTAAGGAGAGAGCCTGGACAGAGAGATCCGAGCGAGTATGAG 60
 QY 135 ValGlyAspArgAspHisArgArgProGlnGluLysLysLysAlaLysGlyLeuGlyLys 154
 Db 61 TT-GGAGACCGAGACCATCGAGGCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 119
 QY 155 GluLeuThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAla 174
 Db 120 GAGATCAGCTCTGCTGATGAGACATCTGAACACGCTGAGTACCCAGAGAGAGAGTGGCT 179
 QY 175 AlaLeuCysLysLysTyrAlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMet 194
 Db 180 GCACCTGTGCAAGAGATGATGCTGAGCTGCTGGAAGAGAGATCGGAACCTCGCAGAGCAGATG 239
 QY 195 LysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGlu 214
 Db 240 AAGCTCTCTCAG 299
 QY 215 HisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArg 234
 Db 300 CACAGCAAGGCTGCTCTGGCCCGAGAGAGAGCTTGAAGCTGTGTGCGGAGAGCTGCAACGG 359
 QY 235 HisAsnArgSerLeuLysGluGlyValGlnArgAlaArgGluGluGluLysArg 254
 Db 360 CACAACCGGCTCCCTGAAG 419

QY 255 LysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuMetGluGln 274
 Db 420 AAGAAGTGCATTTCACATTCAGGTGACACATGAATGACATTCAGTGCATGAACAG 479
 QY 275 HisAsnGluArgAsnSerLysLeuArgGlnGlnAsnMetGluLeuAlaGluArgLeuLys 294
 Db 480 CATAACGAGCGCAATCCCAAGTCGCCAGGAGATATGGAGTAGCCGAGAGCTCAAG 539
 QY 295 LysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAspLysValPheLysHisLys 314
 Db 540 AAGTGTATCGAGCAATACAGCTTCGTGAGGAGCATATCGCAAAAGTCTTCANACATAAG 599
 QY 315 AspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGlnMetLeuLysGlu 334
 Db 600 GACCTGCAGCAGCAGCTAGTGCAGCCCAAGCTCCAGCAGCGCCCGAGGATCTCGAGGAG 659
 QY 335 AlaGluGluArgHisGlnArgGluLysAspPheLeuLeu 347
 Db 660 CGAGAGGAGCGCGACCCAGCGAGAGAGAGTTTCTCTCG 698

RESULT 27

LOCUS BX855901 780 bp mRNA linear EST 16-DEC-2003
 DEFINITION BX855901 AGENAE Rainbow trout multi-tissues subtracted library (tcay) Oncorhynchus mykiss cDNA clone tcay0039b.j.10 5prim, mRNA sequence.

ACCESSION

VERSION BX855901

KEYWORDS

SOURCE EST.

ORGANISM

Oncorhynchus mykiss (rainbow trout)

REFERENCE

1 (bases 1 to 780)
 Govoroun, M., Guiguen, Y. and Le Gac, F.
 Construction and primary characterization of normalized cDNA libraries in rainbow trout, *Oncorhynchus mykiss*
 Unpublished (2003)
 CONTACT: Guiguen Y

JOURNAL

COMMENT

INRA - SCRIBE
 Campus de Beaulieu, RENNES cedex, 35042, France
 Tel: 02.23.48.50.09
 Fax: 02.23.48.50.20
 Email: Yann.Guiguen@beaulieu.rennes.inra.fr
 Sequence cleaned of vector, adaptor and repetitions. Contact us at sigenasupport@jouy.inra.fr to obtain the chromatogram of this sequence.
 Plate: 0039 row: j column: 10
 Seq primer: M13R

FEATURES

source

1..780
 Location/Qualifiers
 /organism="Oncorhynchus mykiss"
 /mol_type="mRNA"
 /db_xref="taxon:8022"
 /clone="tcay0039b.j.10"
 /tissue_type="adipose tissue, blood, brain, differentiating gonads, gills, interrenal, intestine, kidney, liver, muscle, ovary, pituitary, testis"
 /dev_stage="from embryos to adults"
 /lab_host="DH10B"
 /clone_lib="AGENAE Rainbow trout multi-tissues subtracted library (tcay)"
 /note="Vector: pT73D-pac; Rainbow trout multi-tissues - normalized + 1 subtraction (tcay); Clone distribution: AGENAE Resource centre. Francois PIUMI, Francois.Piumi@jouy.inra.fr, INRA, CEA Radiobiologie et Etude du genome (LREG), Domaine de Vilvert, 78352, Jouy-en-Josas cedex, FRANCE"

ORIGIN

Alignment Scores:

Pred. No.: 1.24e-62 Length: 780

Score: 1099.00 Matches: 215
 Percent Similarity: 92.25% Conservative: 23
 Best Local Similarity: 83.33% Mismatches: 20
 Query Match: 40.67% Indels: 0
 DB: 13 Gaps: 0

US-10-023-523-8 (1-530) x BX855901 (1-780)

QY 154 LysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeu 173
 Db 6 GAGGAGATCACCTGCTGATGTCAGACCCCTGAATACCTGAGTACTCCAGAGGATAAAGCTG 65
 QY 174 AlaAlaLeuCysLysLysTyrAlaGluLeuLeuGluGluHisArgAsnSerGlnLysGln 193
 Db 66 ACAGGCTCTGTAAAAAGTATGCTGAACCTCTGGAGGAGCGCGGAAACAACACAGAGCTA 125
 QY 194 MetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGlnLysAspHisLeuArgGly 213
 Db 126 ATCGGGGTCTGCAGAGAAACAGTCTCAGCTGGTCAGGAGACGACCATCTGAGGAAC 185
 QY 214 GluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGln 233
 Db 186 GAACACAGCAAGACCATCTCTGCAGCAGCAAACTGGAGTCGCTCTGTGCGGAGCTGCAG 245
 QY 234 ArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGluGluGluLys 253
 Db 246 AGACACACCGGCACACTCAAGGATGACGGATGACGCGTGCACGTGTGGAGGAGGAGAG 305
 QY 254 ArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGlu 273
 Db 306 AGAAGGAAGTGCAGGCACACTCCAGGTGACTCTGTATGACATCCAGCCAGATGGAG 365
 QY 274 GlnHisAsnGluArgAsnSerLysLeuArgGlnGlnGluAsnMetGluLeuAlaGluArg 293
 Db 366 CAGCATGACGAGAGGAATGCCAGTCTTAGCTGGAGAACTCTGAGCTGCAGAGAAGCTC 425
 QY 294 LysLysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAspLysValPheLysHis 313
 Db 426 AAGAACTCATCCACAGTACGAGTACGAGGAGGAGACATTGACAGGTGTTCAAGCAT 485
 QY 314 LysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGlnMetLeuLys 333
 Db 486 AAGGAAGTGCAGCAACAGTGTGTGATGCTAAGCTACACAGGCTCAGGGTTTGTCTATA 545
 QY 334 GluAlaGluGluArgHisGlnArgGluLysAspPheLeuLysGluAlaValGluSer 353
 Db 546 GAGTCAGAGGAGCCCATCATCAGAGAGAGGAGGACTTTCTGTAAGAGGAGCGGAGAGTCT 605
 QY 354 GlnArgMetCysGluLeuMetLysGlnGlnGlnThrHisLeuLysGlnGlnLeuAlaLeu 373
 Db 606 CAGAGGATGTGTGAGCTCATGAAGCAGCAGGAGGTGCATCTCAACACACAGCTGCTCTTA 665
 QY 374 TyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluValPheThr 393
 Db 666 TACACAGAGAAGTTTGAGGAGTTCAGACGACTCTGTCCACAGCAGCAGAGGTGTTTACC 725
 QY 394 ThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLysLysLysLysLys 411
 Db 726 ACATTCAACAGGAGATGGAGAGATGACAAAGAGATCAAGAAAGCTGGAGAG 779

RESULT 28

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AL637771 702 bp mRNA linear EST 07-NOV-2003
 AL637771 XGC-neurula Silurana tropicalis cDNA clone TNeu020c17 5', mRNA sequence.

AL637771

AL637771.2 GI:38216372

EST.

Silurana tropicalis (western clawed frog)

Silurana tropicalis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;

Xenopodinae; Silurana.

REFERENCE
 1 (bases 1 to 702)
 Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
 Sanger Xenopus tropicalis EST project 2001 (11_2003)
 TITLE
 Unpublished (2003)
 JOURNAL
 On Nov 7, 2001 this sequence version replaced gi:16789750.
 COMMENT
 Contact: Huckle E
 Sanger Institute
 Hinxton, Cambridgeshire, CB10 1SA, UK
 Email: trop@sanger.ac.uk
 This sequence is from a Xenopus Gene Collection (XGC) library
 constructed by Aaron M. Zorn.
 cDNA was oligo dt primed from sug of poly A+ RNA from neurula.
 EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
 5' end and NotI at the 3' end.
 Vector: pCS107; Site 1: EcoRI; Site 2: NotI
 Host: Escherichia coli DH10B
 Sanger Xenopus tropicalis EST project 2001
 TROPICALIS_SEQUENCE_ID: TNeu020c17.pkSP6
 Sequencing primer: SP6
 Location/Qualifiers
 1..702
 /organism="Silurana tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="TNeu020c17"
 /dev_stage="neurula"
 /lab_host="Escherichia coli DH10B"
 /clone_lib="XGC-neurula"
 /notes="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
 was oligo dt primed from sug of poly A+ RNA from neurula.
 EcoRI-NotI cut cDNA was then ligated into pCS107 with
 EcoRI at the 5' end and NotI at the 3' end."
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.67e-61 Length: 702
 Score: 1081.00 Matches: 213
 Percent Similarity: 96.14% Conservativity: 11
 Best Local Similarity: 91.42% Mismatches: 9
 Query Match: 40.01% Indels: 0
 DB: 9 Gaps: 0
 US-10-023-523-8 (1-530) x AL637771 (1-702)
 QY 153 GlyLeuGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLys 172
 DB 2 CGAAGGAGATCACATTATTATGCGAGACTCTGACACACTCAGACCCCGACAGGAAAG 61
 QY 173 LeuAlaAlaLeuCysIleValValAlaGluLeuGluHisArgAsnSerGlnLys 192
 DB 62 CTGACTGCCCTGTGTAAAGATGATCAGAACCTGTGGAGGAGCAGACACATCTCAGAG 121
 QY 193 GlnMetLysLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArg 212
 DB 122 CAAATGAGGACTACTGCAAAAAACACAGCCAGCTTATTCAAGAAAGAGTACTTCGT 181
 QY 213 GlyGluHisSerIleAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLeu 232
 DB 182 ATGAGCAGACAGAGCTATCTCTGCAAGAGCAGCTCGAAAGTTGTTCAGAGAGCTA 241
 QY 233 GlnArgHisAsnArgSerLeuLysGluGluValGlnArgAlaArgGluGluGlu 252
 DB 242 CAGAGACACACCGTACACTAAAGGAGGAGAGTACAGAGGGCTCGTGAGGAGGAG 301
 QY 253 LysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuMet 272
 DB 302 AAACGGCAAGAGTAGTAACTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 351
 QY 273 GluGlnHisAsnGluArgAsnSerLysLeuArgGlnGlnAsnMetGluLeuAlaGluArg 292
 DB 362 CAGCAGCAATGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAA 421
 QY 293 LeuLysLysLeuIleGluGlnThrGluLeuArgGluGluHisIleAspLysValPheLys 312

Db 422 CTTAGAGAGCTTATCGACATATAGCTAAGAGAGACACATTGACAAAGTTTCMAA 481
 QY 313 HisLysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMetLeu 332
 Db 482 CATAAGGATCTTCAACAGCAACTGTGTGATGCAAGCTTCAGCAAGCACAAGAAATGTTG 541
 QY 333 LysGluAlaGluGluArgHisGlnArgGluLysAspPheLeuLeuLysGluAlaValGlu 352
 Db 542 AAAGAGCTGGAGAGCGTCCACCGGGAAGAGATTCTGTGAAAGAGGAGCTCGA 601
 QY 353 SerGlnArgMetCysGluLeuMetLysGlnGlnGlnThrHisLeuLysGlnGlnLeuAla 372
 Db 602 TCCAGCGCATGTGTGAACCTAATGAACAGCAGAGACCCACCTCAACACAGATTAGCA 661
 QY 373 LeuTyThrGluLysPheGluGluPheGlnAsnThrLeu 385
 Db 662 CTGTATACAGAGAGTTTGAGGAGTTTCAAAACACCTTG 700
 RESULT 29
 BU849604 799 bp mRNA linear EST 16-OCT-2002
 LOCUS
 DEFINITION AGENCOURT 10437001 NIH MGC 109 Homo sapiens cDNA clone
 IMAGE:6598234 5', mRNA sequence.
 ACCESSION BU849604
 VERSION BU849604.1 GI:24034567
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 799)
 NIH-MGC http://mgs.nci.nih.gov/
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LCM2830 row: 1 column: 10
 High quality sequence stop: 566.
 Location/Qualifiers
 1..799
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6598234"
 /tissue_type="teratocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_109"
 /note="Organ: ovary; Vector: pOTB7; Site 1: EcoRI; Site 2:
 XhoI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GCCACGAG(G). Library constructed by Ling Hong in the
 Laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC Library."

ORIGIN
 Alignment Scores:
 Pred. No.: 3.07e-61 Length: 799
 Score: 1078.00 Matches: 222
 Percent Similarity: 96.14% Conservativity: 2
 Best Local Similarity: 95.28% Mismatches: 5
 Query Match: 39.90% Indels: 5
 DB: 13 Gaps: 2

US-10-023-523-8 (1-530) x BU618489 (1-799)

QY 1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGlnLysArgProSerGln 20
 Db 89 AAAGACAGCCAGGACACCCAGGAGCAGCCAGGAGCCAGGAGCCAGGAGCCAG 148
 QY 21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProAlaGlyProGlu 40
 Db 149 GCGGCTCTGCGAGTAGAGCAGAGAGTCCCGGAGCAGCAGGCTCTCTCGAAGCCGAG 208
 QY 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 60
 Db 209 GGGGCTCAAGCAGCAGAACCGCTCAGTCTGGGCCCTTCTGATGCTCTGAGGAGCTGAGC 268
 QY 61 ArgGlnLeuGluAlaSerThrTyrCysValAspAsnGlnGlyGlyProGly 80
 Db 269 CGCCAACTGGAGACATACAGCACATCTGTGTGGAGATACACAGGGGGCCCCCGGC 328
 QY 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
 Db 329 GAGGATGGGGCACAGGGTGAGCCGGCTGAACCCGAGATGCGAGAAAGTCCCGGACCTAT 388
 QY 101 ValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGluLysGluProSer 120
 Db 389 GTGGCARGAATGGGAGCCTGACCACTCCAGTCTCAATGGAGAGAGACCTCTC 448
 QY 121 LysGlyAspProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAspHis 140
 Db 449 AAGGGGATCAAAACACAGAGAGATCGGCAGAGTGCAGAGTGCAGAGCCAGACCAT 508
 QY 141 ArgAspProGlnGluLysLysAlaLysGlyLeuGlyLysGluLysLeuMet 160
 Db 509 CGAGGCCACAGGAG 568
 QY 161 GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuLysLysLysTyr 180
 Db 569 CAGACATTTGAATACTCTGAGTACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 628
 QY 181 AlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLys 200
 Db 629 GCTGACTCTGAGGAGCAGCAGAGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 688
 QY 201 GlnSerGlnLeuValGlnGluLysAspHisLeu---ArgGlyGluHisSerLysAla-Va 219
 Db 689 CAGAGCCAGCTGCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 747
 QY 219 LysAlaArgSerLysLeu---GluSerLeuCys 229
 Db 748 CTGGGCCCCGAGCAGAGAGCTTGGAGAGCCCTATGC 782

RESULT 30

BJ618489
 LOCUS
 DEFINITION
 XENOPUS LAEVIS CDNA CLONE XL185E05 5' mRNA SEQUENCE.
 EST.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 XENOPUS LAEVIS (AFRICAN CLAWED FROG)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 Xenopodinae; Xenopus.
 1 (bases 1 to 698)
 Kitayama, A., Terasaka, C., Mochii, M., Ueno, N., Shin-i, T. and
 Kohara, Y.
 EXPRESSED GENES IN X. LAEVIS EMBRYO
 UNPUBLISHED (2001)
 CONTACT: Tadasu Shin-i
 CENTER FOR GENETIC RESOURCE INFORMATION
 NATIONAL INSTITUTE OF GENETICS
 1111 YATA, MISHIMA, SHIZUOKA 411-8540, JAPAN
 TEL: 81-559-51-6856

Fax: 81-559-81-6855

Email: tsin@genes.nig.ac.jp

The information of this clone is available through the following

URL:

http://xenopus.nibb.ac.jp.

FEATURES

source
 Location/Qualifiers
 1..698
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="XL185e05"
 /tissue_type="whole embryo"
 /dev_stage="stage 10.5"
 /clone_lib="NIBB Mochii normalized Xenopus early gastrula library"

ORIGIN

Alignment Scores:
 Pred. No.: 8,75e-61 Length: 698
 Score: 1070.00 Matches: 212
 Percent Similarity: 96.89% Conservative: 6
 Best Local Similarity: 94.22% Mismatches: 7
 Query Match: 39.60% Indels: 0
 DB: 12 Gaps: 0

US-10-023-523-8 (1-530) x BU618489 (1-698)

QY 207 GluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGlu 226
 Db 22 GAGAAGACCACTTCGTAATGAGCATAGCAAGGCCATCTCGCAAGAACCAACTGGAA 81
 QY 227 SerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGlyValGlnArg 246
 Db 82 AGTTTGTGACAGAGCTACAGAGACACACCGTACACTAAAGGAGGAGGATACAGAGG 141
 QY 247 AlaArgGluGluGluLysLysLysLysLysLysLysLysLysLysLysLysLys 266
 Db 142 GCTCGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 201
 QY 267 AspLeuGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLysLysLysLys 286
 Db 202 GACAAACAGTACAGATGGAGCAGACCAATGAACGGAATGCTAAACTTGTCAAGAGAAT 261
 QY 287 MetGluLeuAlaGluArgLysLysLysLysLysLysLysLysLysLysLysLys 306
 Db 262 GTCCGAATCTGTAATCGACTTAAGAAGCTTATTGAGCAATATGAGCTAAGAGAGAGC 321
 QY 307 LysAspLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGln 326
 Db 322 ATTGACAAAGTTTCAAAACATAAGGATCTTCAACACAGCTAGTGACCCAGCTTCAG 381
 QY 327 GlnAlaGlnGluMetLeuLysGluAlaGluGluArgHisGlnArgGluLysAspPheLeu 346
 Db 392 CAAGCACAGAAATGCTGAAAGAGGTGGAAGAGGCTCACCAGCGGGAAGAAATTTCTC 441
 QY 347 LeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHis 366
 Db 442 TTGAAGAGCAGCAGTGGAAATCACAGCCGATGTGCAACTTATGAAACAGCAGAGAGCCAC 501
 QY 367 LeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSer 386
 Db 502 CTCAGCAGCAGCTAGCAGCTGTATACAGAACTTTGAGGAATTTCAAAACACTTTGTTC 561
 QY 387 LysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysIle 406
 Db 562 AAAAGAGTGGAGTTTTCACCACTTTTAAACAAAGAAATGGAAGAGATGCAAGAGAGATC 621
 QY 407 LysLysLeuGluLysGluThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAla 426
 Db 622 AAGAAGCTGGAAAGAGAGACACCATGTATAGTCCCGATGGGAGAGCAGTAAATAGCA 681
 QY 427 LeuLeuGluMetAla 431

Db

682 TTGCTTGTCAAGGCA 696

RESULT 31
BM455349
LOCUS
DEFINITION AGENCOURT_6406997 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:550322
5' mRNA sequence.
ACCESSION BM455349
VERSION BM455349.1 GI:18504389
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: csapbs@mail.nih.gov
Tissue Procurement: Lou Staudt
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Platform: LLML12135 row: b column: 03
High quality sequence start: 90
High quality sequence stop: 669.
Location/Qualifiers
1..1318
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:550322"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 85"
/note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-df primed. Average insert size 1.867 Kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 2,05e-60 Length: 1318
Score: 1069.50 Matches: 256
Percent Similarity: 63.07% Conservative: 19
Best Local Similarity: 58.72% Mismatch: 63
Query Match: 39.58% Indels: 98
DB: 12 Gaps: 9

US-10-023-523-8 (1-530) x BM455349 (1-1318)

QY 178 LysLysTyrrAlaGlueLeuLeuGluHisArgAsn-----Ser 190
Dbb 8 AAACGGGGCGCCGCCACACTTCCTACCCGTAACGAAGGGGGGGGGGAATCTTCG 67
QY 191 GluLysGlnMetLysLeuLeuGlnLysGlnSerGlnLeuValGlnGluLysAspHis 210
Ddb 68 AAAACCGTCGACCACCGCTCGGTACCCCCCAGAGCAGCTGTGTGCAAGAAGAGGACAC 127
QY 211 LeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArg 230
Ddb 128 CTGCGCGGTGAGCACAGAACGGCTCTGCGCCGCGAGCAAGCTTGAGAGCCCTATCCCT 187
QY 231 GluLeuGlnArgHisAsnArgSerLeuLysGluCluLyvAlGlnArgAlaArgGluGlu 250
Ddb 188 CAGTCGACGGCGACAACCGCTCCCTCAAGAAGAAGGTGTGAGCGGGCGGGGAGGAG 247
QY 251 GluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspilleGlnLeu 270

```

LOCUS      BI102887      861 bp      mRNA      linear      EST 26-JUN-2001
DEFINITION 60288485F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:5043572
VERSION     BI102887
ACCESSION  BI102887
KEYWORDS   EST.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 861)
AUTHORS    NIH-MGC http://mgc.nci.nih.gov/
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
            CDNA Library Preparation: Life Technologies, Inc.
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM1119 row: j column: 21
            High quality sequence stop: 738.
FEATURES   Location/Qualifiers
            1..861
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="FVB/N"
            /db_xref="taxon:10090"
            /lab_host="PH108 (T1 phage-resistant)"
            /clone_lib="NCI_CGAP_Kid14"
            /note="Organ: Kidney; Vector: pCMV-SPORT6; Site 1: Not I;
            Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
            Average insert size 1.75 kb. Constructed by Life
            Technologies. Note: this is a NCI_CGAP Library. |"
ORIGIN
Alignment Scores:
Pred. No.:      4,428-60      Length:      861
Score:          1061.00      Matches:    220
Percent Similarity: 88.63%      Conservative: 6
Best Local Similarity: 86.27%      Mismatches: 12
Query Match:     39,27%      Indels:    17
DB:              12          Gaps:       3

US-10-023-523-8 (1-530) x BI102887 (1-861)

QY      288  GluLeuAlaGluArgLeuLysLysLeuLeuGluGlnTyrGluLeuArgGluGluHisLe 307
Db      2   GAGCTAGCGGAGAGCTCAAGAGTTGATCGAGCATACGAGCTTCGTGAGGAGCATATC 61
QY      308  AspLysValPheLysHisLysAspLeuGlnGlnLeuValAspAlaLysLeuGlnGln 327
Db      62  GACAAAGTCTTCAAAACATAAGGACCTGCAGACGACGAGTAGTGACCCCAAGCTCCAGCAG 121
QY      328  AlaGlnGluMetLeuLysGluAlaGluArgHisGlnArgGluLysAspPheLeuLeu 347
Db      122  GCCCAGGAGATCTGAGGAGGACGAGGAGCGGACACCGAGGAGAGGATTTCTCTG 181
QY      348  LysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlnThrHisLeu 367
Db      182  AAGGAAGCGGTGGAGTCCCGAGAGGATGTCGAGCTGATGAGGACGACGAGACCCACCTA 241
QY      368  LysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLysSerLys 387
Db      242  AAGCAGCAGCTCGCCCTGTACAGGAGAGTGTGAGGAGTTCCAGAACACACTTTCCAAA 301
QY      388  SerSerGluValPheThrThrPheLysGlnGluMetGlu-LysMetThrLysLysLeu 407
Db      302  ACCAGTGAAGTGTTCACCACGTTCAACACAGGAGATGGACAGATGACAAAGAGATCAA 361

```

```

QY      407  sLysLeuGluLysGluThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLe 427
Db      362  GAGCTGGAGAAAGAAACACCATGATCGATCCAGGTGGGAAAGACGACACAGGCTCT 421
QY      427  uLeuGluMetAlaGluGluLysThrValArgAspLysGluLeuGlnGlnVal 447
Db      422  GCTGGAGATGGCTGAAGAGAAACCGTCCGGGACAAAGAGCTGGAGGCGCTGCAGGTCAA 481
QY      447  sileGluArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLys 467
Db      482  AATCAGCGGCTGGAGAGCTGTGCCGAGCACTCCAGACTGAGCGCAATGACCTCAACAA 541
QY      467  sArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSerGlyProGluAr 487
Db      542  GAGGATACAGGACCTGACTGCAGGGGC-----ATCATTGACATTGGCTCTGAGCG 592
QY      487  GArgPro-----GluGlyProGlyAlaGlnAl 496
Db      593  GAGGCCAGAGGCCACCACTGCTCCAAAGGAACAGGGGTTGAAAGTCTCTGGGCTCAACC 652
QY      496  aProSerSerProArgValThrGluAlaPcCysTyrProGlyAlaProSerThrGluAl 516
Db      653  AGCCAGCTCTCCAGAGGCCACAGAGCTCTCTGCTGCTCGGAGCCCGGAGCAGCAGGAAC 712
QY      516  aSerGlyGlnThrGlyProGlnGlu-----ProThrSerAla 528
Db      713  AGCAGGCCAGACAGGGCTGGAGAAACCAACCTGTCACTGCC 755

RESULT 33
LOCUS    CD578468      647 bp      mRNA      linear      EST 09-JUL-2003
DEFINITION UI-M-FY0-cfs-n-14-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone
IMAGE:6855471 5', mRNA sequence.
ACCESSION  CD578468
VERSION     CD578468.1 GI:31742859
KEYWORDS   EST.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 647)
AUTHORS    NIH-MGC http://mgc.nci.nih.gov/
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs@mail.nih.gov
            Tissue Procurement: Dr. Jim Lin, University of Iowa
            CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
            CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mousefl.html
            This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
Seq primer: pyX-5.
            Location/Qualifiers
            1..647
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="C57BL/6"
            /db_xref="taxon:10090"
            /clone="IMAGE:6855471"
            /tissue_type="whole brain"
            /dev_stage="embryo 13.5, 14.5, 16.5, 17.5dpc"
            /lab_host="DH10B (T1 phage resistant)"
            /clone_lib="NIH_BMAP_FY0"
            /note="Organ: Brain; Vector: pyX-Asc; Site 1: EcoR I;
            Site 2: Not I; The library was constructed according
            Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
            1996. Denatured RNA was size fractionated on a 1% agarose
            gel. First strand cDNA synthesis was primed with oligo-dT
            primer containing a Not I site. Double strand cDNA was

```


QY 211 uArgGlyGluHisSerLysAlaValLeuAlaAArgSerLysLeuGluSerLeuCysArgGlu 231
 Db 243 GCGAGGGAGACACAGCAAGGCTGTCTGGCCGAGAGAGCTTGAGAGTCTGTGCGCGGA 302
 QY 231 uLeuGlnAArgHisAsnAArgSerLeuLysGluGluValGlnAArgAlaAArgGluGlu 251
 Db 303 GCTGCAACGGCACAAACGGTCTGGAAGGAAGAGCGGTCCAGGAGCGCGTGAGGAGA 362
 QY 251 uGluLysAArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGlu 271
 Db 363 GGAGAACGGCAAGAAAGTACTTACACTTCCAGGTGACACTGAATGACATTCAGTGTGA 422
 QY 271 nMetGluGlnHisAsnGluAArgSerLysLeuAArgGlnGlnAArgMetGluLeuAlaGlu 291
 Db 423 GATGACACACATACAGAGGAACTCCAGCTGGCCGAGAGATATGAGGTAGCCGA 482
 QY 291 uArgLeuLysLeuLeuGluGlnValThrGluLeuAArgGluGluHisIleAspLysValPh 311
 Db 483 GAGGCTCAAGAAGTTGATCGAGCAATACGAGCTTCGTGAGGAGCATATGACAAAGTCTT 542
 QY 311 eLysHisLysAspLeuGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluWe 331
 Db 543 CAAACATPAGACTCTGAGCAGAGCTAGTGGAGCGCCAGCTCCAGAGGCCCGAGGAGAT 602
 QY 331 tLeuLysGluAlaGluAArgHisGlnAArgGluLysAspPheLeu 346
 Db 603 GCTGAGGAGGAGAGAGCGGACAGCCAGGAGAGAGGAGGATTTGTG 648
 RESULT 35
 CAS10893
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 COMMENT
 722 bp mRNA linear EST 15-NOV-2002
 UI-R-FJO-cpv-o-09-0-UI.r1 UI-R-FJO Rattus norvegicus cDNA clone
 UI-R-FJO-cpv-o-09-0-UI 5', mRNA sequence.
 CAS10893
 CAS10893.1 GI:25001847
 EST
 Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 722)
 Bonaldo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 97044477
 889548
 Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 REVERSE
 Location/Qualifiers
 1. .722
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-FJO-cpv-o-09-0-UI"
 /tissue_type="embryo"
 /dev_stage="embryo"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-R-FJO"

FEATURES
 source

/note="Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I;
 UI-R-FJO is a cDNA library containing the following
 tissue(s): rat embryo. The library was constructed
 according to Bonaldo, Lennon and Soares, Genome Research,
 6:791-806, 1996. First strand cDNA synthesis was primed
 with an oligo-dT primer containing a Not I site. Double
 stranded cDNA was ligated to an EcoR I adaptor, digested
 with Not I, and cloned directionally into pVT3-Pac
 vector. The oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is
 located between the Not I site and the (dT)18 tail. The
 sequence tag for this library is CATCTACT. This library
 was created for the University of Iowa Program for Rat
 Gene Discovery and Mapping (Val Sheffield, Bento Soares
 and Tom Casavant)."

ORIGIN

Alignment Scores:
 Pred. No.: 4,346-59 Length: 722
 Score: 1044.50 Matches: 209
 Percent Similarity: 93.48% Conservative: 6
 Best Local Similarity: 90.87% Mismatches: 14
 Query Match: 38.66% Indels: 1
 DB: 14 Gaps: 1
 US-10-023-523-8 (1-530) x CAS10893 (1-722)

QY 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
 Db 3 GAGGAGGAGGAGCTCAGGGGGAGGCCACTGAGCTGAAGACACAGAGAGTCCCGACCTAT 62
 QY 101 ValAlaAArgAsnGlyGluProGluPro---ThrProValValTyrGlyGluLysGluPro 119
 Db 53 CGAGCCAGAGATGGGAGGCTTGACCGGGGCAATCCAGTCTGCAATGGCGAAGAGAGACC 122
 QY 120 SerLysGlyAspProAsnThrGluGluLeuArgLysSerAspGluValGlyAspArgAsp 139
 Db 123 TCTAAGGGGGAGGCTGGCACAGAGAGATCCGAGCAAGTCCAGAGGTTGGAGACCGAGAC 182
 QY 140 HisArgArgProGlnGluLysLysAlaLysGlyLeuGlyLysGluLysThrLeuLeu 159
 Db 183 CTTGAGGCCACAGGAAAGAAAGCAAGGGTCTAGGGAAGGAGATCACTTTGTG 242
 QY 160 MetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaAArgCysLysLys 179
 Db 243 ATGCAGACACTGAACACAGCTGAGCACCCACAGAGAGAGCTGGCTGCACCTGTGCAAGAG 302
 QY 180 TyrAlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLys 199
 Db 303 TACGCCGAGCTGCTGGAGAGCAGCCGCAATGCCAGAGAGCAGCAAGAGCTCTGCAGAGG 362
 QY 200 LysGlnSerGlnLeuValGlnGluLysAspHisLeuAArgGlyGluHisSerLysAlaVal 219
 Db 363 AAGCAGAGCCAGCTGTGTCAGAGAGAGGACCATCTGGTGTGTGAGCACAGGAGGCTGTC 422
 QY 220 LeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnAArgHisAsnArgSerLeu 239
 Db 423 CTGGCCCGAAGTAAGTGTAGAGCTGTGCGGGAGCTGCAACCGGCACACCGATCCCTA 482
 QY 240 LysGluGluGlyValGlnAArgAlaArgGluGluLysArgLysGluValThrSer 259
 Db 483 AAGGAAGAGGTGTGACGAGGCCCGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 542
 QY 260 HisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsn 279
 Db 543 CACTTCAGAGGTGACGTGAATGACATTCAGCTGCAATGGAGGAGCATATGAGCGGAGAC 602
 QY 280 SerLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuLeuGln 299
 Db 603 GCCAAGCTGCGACTAGAGAACTGAGCTGGCCGAGAGGCTGCAAGAGGAGGAGGAGGAG 662
 QY 300 TyrGluLeuArgGluGluHisIleAspLys 309

```

Db      663 TAGCAGCTTCGTGAGGACGAGTTTTCAAA 692

RESULT 36
AL887857
LOCUS      665 bp mRNA linear EST 04-DEC-2003
DEFINITION AL887857 XGC-egg Silurana tropicalis cDNA clone TEG9080c08 5', mRNA
sequence.
ACCESSION AL887857
VERSION    AL887857.2 GI:38689228
KEYWORDS   EST.
SOURCE     Silurana tropicalis (western clawed frog)
ORGANISM   Silurana tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Silurana.
1 (bases 1 to 665)
Croning M.D.R., Ashurst J.L., Taylor R., Zorn A.M. and Rogers, J.
Sanger Xenopus tropicalis EST project 2001 (11_2003)
Unpublished (2003)
On Sep 16, 2002 this sequence version replaced gi:22938408.
Contact: Taylor R
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: tropesanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: TEG9080c08.plkSP6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
cDNA was oligo dt primed from Sug of poly A+ RNA from egg.
ECORI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
5' end and NotI at the 3' end.
Vector: pCS107; Site_1: EcoRI; Site_2: NotI
Host: Escherichia coli XLI-blue.
Host: Escherichia coli XLI-blue.
Location/Qualifiers
1. .665
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TEG9080c08"
/dev_stage="egg"
/lab_host="Escherichia coli XLI-blue"
/clone_lib="XGC-egg"
/notes="vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
was oligo dt primed from Sug of poly A+ RNA from egg.
ECORI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end"

FEATURES
source
1. .665
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TEG9080c08"
/dev_stage="egg"
/lab_host="Escherichia coli XLI-blue"
/clone_lib="XGC-egg"
/notes="vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
was oligo dt primed from Sug of poly A+ RNA from egg.
ECORI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end"

ORIGIN
Alignment Scores:
Pred. No.: 7,75e-59 Length: 665
Score: 1040.00 Matches: 205
Percent Similarity: 97.29% Conservative: 10
Best Local Similarity: 92.76% Mismatches: 6
Query Match: 38.49% Indels: 0
DB: 9 Gaps: 0

US-10-023-523-8 (1-530) x AL887857 (1-665)

Qy      152 LeuGlyLysGluLeuThrLeuMetGlnThrLeuAsnThrLeuSerThrProGluGlu 171
Db      1 CTGGGAAGGAGATCACATTTATTCAGACTCTGACACACTCAGCAGCCCGAGAGAA 60

Qy      172 LysLeuAlaLeuLysLysLysAlaGluLeuGluGluHisArgAsnSerGln 191
Db      61 AAGCTGACTGCCTGTGAAGAGTATGACAACTGTTGGAGGAGCAGACACTCTCAG 120

Qy      192 LysGlnMetLysLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeu 211
Db      121 AAGCAATGAGGACTACTGCAAAAAAAGCAGCCAGCTTATTCAGAAAGGAGTCTACTT 180

Qy      212 ArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLysLeuGluSerLeuCysArgGlu 231

```

```

Db      181 CGTAATGACACACCAAGGCTATCTCGCAAGCAAGCAAGCTCGAAAGCTTTGTGCAGAG 240
Qy      232 LeuGlnArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGluGlu 251
Db      241 CTACAGACACACCAAGGCTATCTCGCAAGCAAGCAAGCTCGAAAGCTTTGTGCAGAG 300
Qy      252 GluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGln 271
Db      301 GAGAAACGCAAGGATTAACCTCCATTTCCAGTGACACTCAATGACATCCAGTCACAG 360
Qy      272 MetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGlnGlnAsnMetGluLeuAlaGlu 291
Db      361 ATGCAGCAGCACAATGAACGTAATGCTAAACTGCTCAAGAGAATGTTGAGCTTGTCTGAT 420
Qy      292 ArgLeuLysLysLeuLysGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 311
Db      421 CGACTTAAGAGCTTATCGAGCAATATGAGCTAAGAGAAGCAGCAGCATTGACAAAGTTTC 480
Qy      312 LysHisLysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnGlnGlnMet 331
Db      481 AAACATAAGGATCTTCACACCAACTGCTGATCGAAGCTTCAGCAAGCACAAGAAATG 540
Qy      332 LeuLysGluAlaGluGluArgHisGlnArgGlnLysAspPheLeuLysGluAlaVal 351
Db      541 CTGAAGAAGTGGAGAGCGCTCACCGCGGAAAGAGGTTTCTGTGAAGAGGCGAGTC 600
Qy      352 GluSerGlnArgMetCysGluLeuMetLysGlnGlnGlnGlnGlnGlnGlnGlnGln 371
Db      601 GATCCCCAGCGCAATGTGTGAACCTAATGAAACAGCAGAGACCCACCTCAACACAGTTA 660
Qy      372 Ala 372
Db      661 GCA 663

RESULT 37
BM963869
LOCUS      735 bp mRNA linear EST 18-MAR-2002
DEFINITION UI-M-EQO-bwn-b-08-0-UI.r1 NIH_BMAP_EQO Mus musculus cDNA clone
IMAGE:5698471 5', mRNA sequence.
ACCESSION BM963869
VERSION    BM963869.1 GI:19547289
KEYWORDS   EST.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 735)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
cDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
Location/Qualifiers
1. .735
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5698471"
/tissue_type="whole brain"
/dev_stage="embryo 15.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_EQO"

```

/note="Organ: brain; Vector: pYX-Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is CAGCAGCAGC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:
Pred. No.: 8,76e-59 Length: 735
Score: 1040.00 Matches: 208
Percent Similarity: 89.50% Conservative: 5
Best Local Similarity: 87.39% Mismatches: 23
Query Match: 38.49% Indels: 2
DB: 12 Gaps: 2

US-10-023-523-8 (1-530) x BM963869 (1-735)

Qy 1 LysSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 20
Db 25 AAGGCGACCGGGGACCGGAGCAGGACCGGAGGAGCCGATCGGAGCAGCAG 84
Qy 21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40
Db 85 ACAGCTCTCTGGGCGGAGCGGAGGT--TCCACAGCGCGGCTCTGGGAGACCGG 141
Qy 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 60
Db 142 GGGCTCGAGCTAAGACAGCTCAGCTGGGCGCTCTGTGAGCTCTCAGGAGCTGAGC 201
Qy 61 ArgGlnLeuGluAspLeuSerThrTyrCysValAspAsnGlnGlyProGly 80
Db 202 CGGCAGTGAAGACATCTCTGAGTACATCTGTGTGACAAATCAGGAGGCGCGCT 261
Qy 81 GluAspGlyAlaGlnGlyProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
Db 262 GAGGAGGAGCAGCGGTGAGCCACTGAGCGGAGACACGAGAGTCCCGACCTAT 321
Qy 101 ValAlaArgGlnGlyProGluPro---ThrProValValTyrGlyGluLysGluPro 119
Db 322 GCAGCCAGGAGTGGGAGCGCTGAACCGAGCATTCCTCAGTCTGCAACGCGGAGAGAG 381
Qy 120 SerLysGlyAspProGlnThrGluGluLeuLeuArgGlnSerAspGluValGlyAspArgAsp 139
Db 382 TCTAAGGAGAGCTGGAACAGAGGAGATCCGAGCGAGTGATGAAGTTGGAGCCGAGC 441
Qy 140 HisArgProGlnGluLysLysAlaLysGlyLeuGlyLysGluLeuThrLeuLeu 159
Db 442 CATCGGAGCCACAGGAGAGAGAAAGCCAGGGTCTAGGGAAGGAGATCACTCTGCTG 501
Qy 160 MetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLys 179
Db 502 ATCGAGACATGACACCGCTAGTATCCAGAGGAGAGCTGGCTGCCTGTCAGAGAG 561
Qy 180 TyrAlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLys 199
Db 562 TATGCTGAGCTGTGGAAGAGCATCGGAATCGCAGAGCAGATGAAGCTCTCGAGAG 621
Qy 200 LysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaVal 219
Db 622 AAGCAGACCGAGCTCTGTGAGAGAGAGGACCATCTCCGAGGGGAGACACAGCAGGCTGTC 681
Qy 220 LeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArg 237

Db 682 CTGGCCGGAAGCAAGCTTGAGAGTCTGTGCCGGAGCTGCAACGCGCACACCGG 735
RESULT 38
BU231392
LOCUS
DEFINITION
603947670F1 CSEQCHN23 Gallus gallus cDNA clone ChEST902117 5', mRNA
sequences.
BU231392 782 bp mRNA linear EST 26-NOV-2002
BU231392.1 GI:25473103
VERSION
KEYWORDS
SOURCE
ORGANISM
Gallus gallus (chicken)
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE
1 (bases 1 to 782)
Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
22335534
12445392
Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES
source

1..782
Location/Qualifiers
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="ChEST902117"
/dev_stage="22"
/lab_host="DH10B"
/clone_lib="CSEQCHN23"
/note="Organ: heads; Vector: pBluescript II KS(+); Site_1:
EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

ORIGIN

Alignment Scores:
Pred. No.: 1.49e-58 Length: 782
Score: 1037.00 Matches: 193
Percent Similarity: 91.51% Conservative: 44
Best Local Similarity: 74.52% Mismatches: 22
Query Match: 38.38% Indels: 0
DB: 13 Gaps: 0

US-10-023-523-8 (1-530) x BU231392 (1-782)

Qy 199 LysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAla 218
Db 5 AAAAAGCAAGCACAAGTCTGTGAAGGAGAAAGTCCACTTGCAGAGCGAGCATAGCAAGGCC 64
Qy 219 ValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSer 238
Db 65 ATTTTGGCAGCGCAACTGGAATCTCTCTGTCGGAGCTTCAGGCTCATACCAAGACT 124

426 AlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGluLeuGluGlyLeuGln 445
 722 GCCTACTGCAATGCTGAGAGAACTGACGACACAGAAATACAAAGCTTGGAG 781
 446 ValLysLeuGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsn 463
 782 GTTAAATAGAGGCTGTAGAGAACTGTCAGGCACTTCAACCGAGAGAAAC 835

RESULT 40
 BQ389733 654 bp mRNA linear EST 22-MAY-2002
 LOCUS NISC_mg09c06.y1 NICHED_XGC_Emb5 Silurana tropicalis cDNA clone
 DEFINITION IMAGE:5308018.5, mRNA sequence.
 BQ389733
 VERSION BQ389733.1 GI:21077420
 KEYWORDS EST.
 SOURCE Silurana tropicalis (western clawed frog)
 ORGANISM Silurana tropicalis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 Xenopodinae; Silurana.
 1 (bases 1 to 654)
 NIH-XGC http://image.llnl.gov/image/html/xenopuslib info.shtml.
 National Institute of Child Health and Human Development, National
 Cancer Institute, Xenopus Gene Collection
 Unpublished (2002)
 CONTACT: Robert Strausberg, Ph.D.
 EMAIL: cgaabs-remail.nih.gov
 CDNA Library Preparation:
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
 Sequencing Center (NISC)
 Cloned unidirectionally. Primer: Oligo dT. Average insert
 size 2.0 kb. Constructed by Invitrogen. Note: This is a
 Xenopus Gene Collection (XGC) library."

Alignment Scores:
 Pred. No.: 1-88e-58 Length: 654
 Score: 1034.00 Matches: 204
 Percent Similarity: 98.16% Conservative: 9
 Best Local Similarity: 94.01% Mismatches: 4
 Query Match: 38.27% Indels: 0
 DB: 13 Gaps: 0

US-10-023-523-8 (1-530) x BQ389733 (1-654)

199 LysLeuGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerIysala 218
 3 AAAAACAAGACCCAGCTTATCAAGAAAGGATCTCTCGTAATGACACAGCAGCA 62
 219 ValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSer 238
 63 ATCTGCGCAAGCAAGCTGCAAGCTTGTGACAGAGCTACAGAGACACAAACCGTACA 122
 239 LeuLysGluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThr 258

123 CTAAGAGGAGGAGTACAGAGGCTCTGAGGAGGAGAGAAACGCAAGACTRACC 182
 259 SerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGluHisAsnGluArg 278
 183 TCCCATTTTCAAGTGCAGCTCAATGACATCCAGTACAGATGCGACGACCAATGACGT 242
 279 AsnSerLysLeuArgGlnGlnAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGlu 298
 243 AATGCTAAACTCGTCAAGAGAAATGTTGAGCTTGTGATCGACTTAAGAAGCTTATCGAG 302
 299 GlnTyrGluLeuArgGlnGluHisIleAspLysValPheLysHisIysAspLeuGlnGln 318
 303 CAATATGAGCTAAGAGAGAGACATGTGAAAGATTTTCAACATAGAGATCTTCAACAG 362
 319 GlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluGluArg 338
 363 CAATGCTGATGCGAGGCTTCAGCAAGCACAAGAAATGCTGAAGAAGTGAAGAGCGT 422
 339 HisGlnArgGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGlu 358
 423 CACCACGCGGAAAAGGAGTTTCTGTGAAGAGGAGTGAATCCCGCATGTGTGAA 482
 359 LeuMetLysGlnGlnGluThrHisLysLysGlnGlnLeuAlaLeuTyrThrGluLysPhe 378
 483 CTATGCAACACAGACAGACACCCACCTCAACACACAGTAGCTGTATACAGAGAGTTT 542
 379 GluGluPheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGlu 398
 543 GAGGAGTTTCAAAACACCTTGTCTAAAGCAGTGAATCCCGCATGTGTGAA 602
 399 MetGluLysMetThrLysLysLysLysLysLysLysLysLysLysLysLysLysLys 415
 603 ATGGAAGAAGATGACAAAGAGATTGAAGAGCTGGAAGAAGAGACACCAATG 653

RESULT 41
 BQ950691 930 bp mRNA linear EST 21-AUG-2002
 LOCUS AGENCOURT_8752844 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6335682
 DEFINITION 5', mRNA sequence.
 BQ950691
 VERSION BQ950691.1 GI:22366169
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 930)
 NIH-MGC http://mgi.mgi.nhl.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 EMAIL: cgaabs-remail.nih.gov
 Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman,
 Ph.D.
 CDNA Library Preparation: ResGen, Invitrogen Corp
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Cloned unidirectionally. Primer: Oligo dT. Average insert
 size 2.0 kb. Constructed by Invitrogen. Note: This is a
 Xenopus Gene Collection (XGC) library."

Alignment Scores:
 Pred. No.: 1-88e-58 Length: 930
 Score: 1034.00 Matches: 204
 Percent Similarity: 98.16% Conservative: 9
 Best Local Similarity: 94.01% Mismatches: 4
 Query Match: 38.27% Indels: 0
 DB: 13 Gaps: 0

US-10-023-523-8 (1-530) x BQ950691 (1-930)

199 LysLeuGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerIysala 218
 3 AAAAACAAGACCCAGCTTATCAAGAAAGGATCTCTCGTAATGACACAGCAGCA 62
 219 ValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSer 238
 63 ATCTGCGCAAGCAAGCTGCAAGCTTGTGACAGAGCTACAGAGACACAAACCGTACA 122
 239 LeuLysGluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThr 258

Oligo dt. Average insert size 1.95 kb. Constructed by ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ORIGIN
Alignment Scores:
Pred. No.: 1,13e-57 Length: 930
Score: 1025.00 Matches: 218
Percent Similarity: 84.70% Conservatives: 9
Best Local Similarity: 81.34% Mismatches: 34
Query Match: 37.93% Indels: 7
DB: 13 Gaps: 3

US-10-023-523-8 (1-530) x BQ950691 (1-930)

QY 1 LysSerSerProGlyGlnProGluAlaGlyProGlyGlnAlaGlyProGluGlyAlaGlnGluArgProSerGln 20
DB 96 AAGGGGAGCGCGGGGCAACGGGAAGCAGGAGCGCGGAGGAGCGCCATGTCAGCAGCCAGCAG 155

QY 21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40
DB 156 ACAGCTCTCTGGGCGGAGACCGGAGGT---TCCACAGCCAGGCTCTCTGGGAGACCGAG 212

QY 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 60
DB 213 GGGGCTCGAGCTAAGCAGCTCAGCTGGGGCCCTCTGTGAGCTCTCTGAGGAGCTGAGC 272

QY 61 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnGlnGlyGlyProGly 80
DB 273 CGGAGTTGGAGACATCTGTAGTACATCTGTGTGACAAACATCAGGGAGGCGCGCT 332

QY 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
DB 333 GAGGAGGAGCAGCGTGTGAGCCACTGAGCGCGGAGACACGGAGAGTCCGACCTAT 392

QY 101 ValAlaArgAsnGlyGluProGluPro---ThrProValValTyrGlyGluLysGluPro 119
DB 393 CGAGCCAGGAAATGGGAGGCTGAACCGAGCATTCAGTCCGTCAGTCCGTCAGTCCGTCG 452

QY 120 SerLysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAsp 139
DB 453 TCTAGGAGAGCCTGGACAGAGAGATCCGAGCGAGTGTAGTGTGGAGACCGAGAC 512

QY 140 HisArgArgProGlnGlyLysLysLysLysLysLysLysLysLysLysLysLysLys 159
DB 513 CATCGAGGCCACAGGAG 572

QY 160 MetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaAlaLeuCysLysLys 179
DB 573 ATGCAGACACTGACACCGCTGATGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 632

QY 180 TyrAlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLys 199
DB 633 TATGTGAGTCTGTGAAGAGCATCGGAACATCGGAGAGAGAGAGAGAGAGAGAGAGAGAG 692

QY 200 LysGlnSerGlnLeuValGlnGlnLysAspHisLeuArgGlyGluHisSerLysAlaVal 219
DB 693 AAGCAGAGCCAGCTGTGTGAG 752

QY 220 LeuAlaArgSerLysLeuGluSerLeuCysArg-GluLeuGlnArgHisAsnArg-SerL 239
DB 753 CTGGCCCGAGCAAGCTTGAGAGTGTGTGCGGGAGAGTCCACGGGAGAGAGAGAGAGAGAG 812

QY 239 euLysGluGluGlyVal-----GlnArgAlaArgGluGluGluGlyArgLysGluV 257
DB 813 TGAAGAGAGAGGCGTCCCGGAG 872

QY 257 alThr-SerHisPheGln 262
DB 873 GAACCTTCCCACTTCCAG 890

RESULT 42

BU151934

LOCUS

BU151934

954 bp

mRNA

linear

EST 03-SEP-2002

DEFINITION

AGENCOURT 8753209 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6394355 5', mRNA sequence.

ACCESSION

BU151934

VERSION

BU151934.1 GI:22665466

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 954)

AUTHORS

NTH-MGC <http://mgs.nci.nih.gov/>.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgaps-remail.nih.gov

Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman, Ph.D.

cDNA Library Preparation: ResGen, Invitrogen Corp

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L14M13887 row: e column: 12

High quality sequence stop: 512.

Location/Qualifiers

1. .954

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE:6394355"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_130"

/notes="Organ: otocysts; Vector: pCMV-SPORT6.1; Site: 1: EcoRV; Site: 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.95 kb. Constructed by ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

395 PheLysGlnGluMetGluLysMetThrLysLysLysLysLysLysGluThrThr 414
 361 TTTCAACAGGAGATGGAAGATGCAACAAAGATCAAGAGCTGGAGAAAGAACACC 420
 415 MetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeuMetAlaGluGluLys 434
 421 ATGTATCGATCCCGTGGGAAACAGCAACAGCTCTGCTGGAGATGCTGAAGAGAAA 480
 435 ThrValArgAspLysGluLeuGluGluGluValLysLysLysLysLysLysLeu 454
 481 ACCTC-CGGACAA-GAGCTGGAGGCGCTGCAGGTGAATC-CAGCGGCTGGAGAGCTG 537
 455 CysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAla 474
 538 TGGGAGCATGAGACCGAGCGCATGACCTCAACAGAGGTACAGACCTGACTGCA 597
 475 GlyGlyGlnGlySerLeuThrAspSerGlyProGluArgArgProGlu----- 490
 598 GGGGGC-----ATCACTGACATTGGCTCTGAGCGGAG-CCAGAGGCCACACTGCCT 647
 491 -----GlyProGlyAlaGlnAlaProSerSerProArgValThrGlu 504
 648 CAAGGAACAGGGGNNTGAACTCTGGGGCTCACCC-AGCAGCTCTCCNAGGGC-ACAGAC 705
 505 AlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGlu 524
 706 GCTCTTCTGCTCTGCGAGGCGCCGAGCACAGGAACAGCAGCGCCCTGGAGAG 765
 525 ProThrSerAlaArgAla 530
 766 CCCACCCCTGCACTGCC 783

RESULT 43

BF796069
 LOCUS 602258906F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4342392 5',
 DEFINITION mRNA sequence.

ACCESSION BF796069

VERSION BF796069.1 GI:12101123

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Louis Staudt, M.D., Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLAM9957 row: k column: 01

High quality sequence stop: 557.

FEATURES

source

1..965
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4342392"
 /tissue_type="lymphoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 85"
 /note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.867 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores: 1.89e-55. Length: 965
 Pred. No.: 991.50 Matches: 227

Score: 82.94% Conservatives: 16

Percent Similarity: 77.47% Mismatches: 28

Best Local Similarity: 36.70% Indels: 25

Query Match: 10 Gaps: 2

DB: 1.0

US-10-023-523-8 (1-530) x BF796069 (1-965)

QY 259 SerHisPheGlnValThrLeuAsnAspLleGlnLeuGlnMetGluGlnHisAsnGluArg 278
 Db 1 TCGAACTTCCAGTCCACACTGAATGCATTCAGCTGCAGATGGAAACACACCAATGAGCGC 60
 QY 279 AnSerLysLeuArgGlnGlnuAsnMetGluLeuAlaGluArgLeuLysLysLeuLeu 298
 Db 61 AACTCAAGCTCGCGCAAGAGACATGAGCTGGCTGAGAGGCTCAAGAAGCTGATTGAG 120
 QY 299 GlnTyrGluLeuArgGluGluHisLysPheLysValPheLysHisLysAspLeuGln 318
 Db 121 CAGTATGAGCTCGCGGAGGAGCATATCGACAAGTCTTCAACACACAGGAGCTACACAG 180
 QY 319 GlnLeuValAspAlaLysLeuGlnGlnAlaGlnGlnMetLeuLysGluAlaGluArg 338
 Db 181 CAGCTGGTGGATGCCAAGCTCCAGCAGGCCAGGAGATGCTAAAGAGGCGCAGAGCGG 240
 QY 339 HisGlnArgGluLys-AspPheLeuLysGluAlaValGluSerGlnArgMetCysG 358
 Db 241 CACCAGCGGAGAGGCGCATTTCTCTCTGAAAGAGGAGTAGAGTCCAGAGGATGTGTA 300
 QY 358 uLeuMetLysGlnGlnGluThrHisLysLysGlnGlnLeuAlaLeuTyrThrGluLysP 378
 Db 301 GCTGTATGAAGCAGCAGAGAGACCCAGCTGAAGCAACAGCTTGCCCTATACACAGAG 360
 QY 378 eGluGluPheGlnAsn-ThrLeuSerLysSerSerGluValPheThrThrPheLysG 398
 Db 361 TGAGGAGTTCCAGAACACACTTCCAAAGAGCAGCGAGGTATTCCACCATTCAGAGCG 420
 QY 398 luMetGluLysMetThrLysLys-1leLysLysLeuGlu-LysGluThrThrMetTyr 417
 Db 421 AGATGAAAGATGACTAAGAGCATCAAGAGCTGAGCAAGAGAAACACCATGTACCG 480
 QY 417 gSerArgTyrGluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluThrValAr 437
 Db 481 GTCCCGTGGGAGAGCAGCAACAGGCCCTGCTTGAGATGGCTGAGGG-GAAACAGTCCG 539
 QY 437 gAspLysGluLeuGluGluGlnValLysLysLysGlnArgLeuGluLysLeuCysArg 457
 Db 540 GGATACAGAACTGGAGGGCTGCAGGCTCAACAATCCACCGGTGGAGAGCTGTCCGCGC 599
 QY 457 aLeuGlnThrGluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGly 477
 Db 600 CACTGAAGACAGAGCGCATGACCTGAACAGAGGGGTACG-GACCTGAGTGGCGGTGGCCA 558
 QY 477 nGlySerLeuThrAspSerGlyProGluArgArgProGlu-GlyProGlyAlaGlnAla 497
 Db 659 GGGACACCTCACTGACGT-GGCCCTGAGAGAGGCCAGAGGGGGCTCTGGGGCCAG--C 714
 QY 497 roSerSerProArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAla 517
 Db 715 CCCCAGGCCGCCAGGGGTACACGAAGCGGCTGTCTACCCAGGCGCACCGAGAGCTGTG 774
 QY 517 erGlyGln-----ThrG 521
 Db 775 CAGGCGGAGGGGCTTACGAGCCCTCCGCCAGGCTTCAACCTGAGGCGGAGGAGCG 834
 QY 521 lyProGlnGluProThrSerAlaArgAla 530
 Db 835 GAGCAGCGCGCACCAAGAGCGCGCGAGGCA 863

RESULT 44

AY405165 1611 bp DNA linear GSS 12-DEC-2003
 LOCUS Homo sapiens HCM2133 gene, VIRTUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.
 DEFINITION
 AY405165
 VERSION AY405165.1 GI:39761139
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1611)
 AUTHORS Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
 Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
 Adams, M.D. and Cargill, M.
 TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302
 REFERENCE 2 (bases 1 to 1611)
 AUTHORS Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
 Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
 Adams, M.D. and Cargill, M.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 source
 1..1611
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 gene
 <1..>1611
 /locus_tag="HCM2133"
 ORIGIN
 Alignment Scores:
 Pred. No.: 5,99e-55 Length: 1611
 Score: 988.00 Matches: 206
 Percent Similarity: 61.07% Conservative: 56
 Best Local Similarity: 48.02% Mismatches: 107
 Query Match: 36.57% Indels: 60
 DB: 29 Gaps: 2
 US-10-023-523-8 (1-530) x AY405165 (1-1611)
 QY 160 MetGlnThrLeuAenThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLys 179
 Db 1 ATGCAAAATCTGAACAAGTTGCAACACCGGAGAAAAGTTTATTTTATTCAGAAG 60
 QY 180 TyrAlaGluLeuLeuGluHisArgAsnSerClnLysGlnMetLysLeuLeuGlnLys 199
 Db 61 TATCTGAATTCGTGGATGAACATCGTACTGAGCAAAAGAAAGTTAAAGTCTCTCCAAAG 120
 QY 200 LysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaVal 219
 Db 121 AAACAGGTACAAATTCAAAAGAAAGACCCAGTTACAGGTGAACACACAGAGTATC 180
 QY 220 LeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeu 239
 Db 181 CTCCTCGAAGCAATTTGGAGAGTCTGTCCGGAGCTGCAGAGACACACACAGACTCTG 240
 QY 240 LysGluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSer 259
 Db 241 AAGGAAGAGCGCTTCAGCGGCACGTCGAGAGAGAGAGAGAAAGGAAGGAATACACAGC 300
 QY 260 HisPheGlnValThrLeuAenAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsn 279
 Db 301 CATTTCCAGAGTACCCCTCACGGACATCCAGGGCCAGATCGAGCAGCAGAGTGGCGAAAT 360

QY 280 SerLysLeuArgGlnGlnAenMetGluLeuAlaGluArgLeuLysLysLeuIleGluGln 299
 Db 361 ATGAAGCTCTGTCCAGGAAACACAGAGCTGCAGAAAAGCTGAAAGCATCATCGATCAG 420
 QY 300 TyrGluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGln 319
 Db 421 TATGAGCTCAGAGGAGCATCTGGACAAAATATTATAACACAGAGAACTGCGACGAAG 480
 QY 320 LeuValAspAlaLysLeuGlnGlnAlaGlnMetLeuLysGluAlaGluGlnArgHis 339
 Db 481 CTGGTGCATGCAAGCTTCAGCAGCGCCCAAGAAATGATGAAGGAGCGGAGGCGACAC 540
 QY 340 GlnArgGluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeu 359
 Db 541 AAACGAGAAAGGAATATTGTCGAACCCAGGAGCAGAGTGGAACACTTCAGCGGAAGTG 600
 QY 360 MetLysGlnGlnGluThrHisLysLysGlnGlnLeuAlaLeuTyrThrGluLysPheGlu 379
 Db 601 CTGAGGAGCAGACAGACAGTCTCTGAGGCTCAGCTCACTCTCTACTCAGAAAGTTTGA 660
 QY 380 GluPheGlnAenThrLeuSerLysSerGluValPheThrThrPheLysGlnGluMet 399
 Db 661 GAATTCAGAGCAGACACTAACTAAAGCAACGAGGTGTTTGCACCTTCAACACGGAATG 720
 QY 400 GluLysMetThrLysLysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArg 419
 Db 721 GACAAANNN 780
 QY 420 TrpGluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLys 439
 Db 781 NNN 840
 QY 440 GluLeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuLysCysArgAlaLeuGln 459
 Db 841 GAATATGAGTGTGTTGTGATGAAATCGGAGGCTAGAGAACCTCTGCGGTGCTTTACAA 900
 QY 460 ThrGluArgAsnAspLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySer 479
 Db 901 GAAGAGAGAAACGAATCCACAAAATAATCAGAGCGCAGAGAAATATCTGAAAAGGATGAC 960
 QY 480 LeuThrAspSerGlyProGluArgArgProGluGly----- 491
 Db 961 CAAAGTCAGACACAACCTCCGATGAGAGCGCAGAGTCAAAACGTCTCTGTGGATCAAGAGATT 1020
 QY 491 ----- 491
 Db 1021 GAGCAGAGGAGTTAATAGTGTCCAAACCGCGTGAATAATCTGGCCACAGCCTTCATG 1080
 QY 491 ----- 491
 Db 1081 ATAATTCATCCAGAGTCAACCCCGCAGCTCCAAAGAAACCCCAACCCGAAATAGGC 1140
 QY 492 -----ProGlyAlaGlnAlaProSerSerPro----- 500
 Db 1141 AGTCTCAGAGAGTGTGAGCGCGCTCTCAGAGAGCGCAGCAACCCCTCTCATCCT 1200
 QY 501 ---ArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGln 519
 Db 1201 TCACGGGATTTCAGAGAGTCCCTGCTCCCTAACTCTCAGGCTGAAGCGGAGGAGGC 1260
 QY 520 ThrGlyProGlnGluProThrSerAla 528
 Db 1261 AGTGTGTGAACCTCCCTCCAGGCC 1287
 RESULT 45
 AW412004 620 bp mRNA linear EST 08-FEB-2000
 LOCUS u055f09.y1 NCI CGAP Lu29 Mus musculus cDNA clone IMAGE:264673 5'
 DEFINITION similar to SW:IL14_HUMAN F40222 INTERLEUKIN-14 PRECURSOR ;, mRNA
 sequence.
 ACCESSION AW412004
 VERSION AW412004.1 GI:6937859
 KEYWORDS EST.

255 Lys-GluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluG1 274
 538 AAGCAGGTGACATCCCACTTCAGGTGACACTGAATGACATCCAGCTCCAGATGGAGCA 597
 274 nHisGluArgAsnSerLysLeuArgGlnGlnAsnMetGluLeuAlaGluArgLeuLy 294
 598 GCACACGAGAGAACTCAAGCTGCGCAGGAGAACTGGAGCTGGCAGCGGCTCA 657
 294 sLysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAspLys-ValPheLysHisL 314
 658 GAGCTCATTGACAGTATGAGTACGGGAGGAGCACCATTGATGAGGGTGTTCACACACA 717
 314 YS 314
 718 AG 719

RESULT 48
 BQ421247
 LOCUS
 DEFINITION BQ421247 813 bp mRNA linear EST 23-MAY-2002
 AGENCOURT 7827164 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6014257
 5', mRNA sequence.
 ACCESSION BQ421247
 VERSION BQ421247.1 GI:21116562
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 813)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-rc@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1AM13208 row: h column: 02
 High quality sequence stop: 660.

FEATURES
 source
 1..813
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6014257"
 /tissue types="embryonal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_92"
 /notes="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 2.5 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

ORIGIN
 Alignment Scores:
 Pred. No.: 2.17e-54 Length: 813
 Score: 974.00 Matches: 190
 Percent Similarity: 89.15% Conservative: 40
 Best Local Similarity: 73.64% Mismatches: 28
 Query Match: 36.05% Indels: 1
 DB: 13 Gaps: 0

US-10-023-523-8 (1-530) x BQ421247 (1-813)

146 LysLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThr 165
 18 AGGACAAAGAGAAACCTTAGGAAAGAAAGCTTTTATTACTGATGCAAGCCCTTAACACC 77

166 LeuSerThrProGluGluLysLeuAlaLeuCysLysLysTyrAlaGluLeuGlu 185
 78 CTTTCAACCCAGAGGAGAGCTGGCAGCTCTCTGTAGAAATATGCTGATCTTCTGAG 137
 186 GluHisArgAsnSerGlnLysGlnMetLysLeuGlnLysLysGlnSerGlnLeuVal 205
 138 GAGAGCAGGAGTGTTCAGAGCAATGAAGATCTCTGCAGAAAGCAAGCCAGATTGTG 197
 206 GlnGluLysAspHisLeuArgGlyGluHisSerLysLysAlaValLeuAlaAArgSerLysLeu 225
 198 AAGAGAAAGTTCATTTCAGAGTGAACATAGCAGGCTATCTTGGCAAGAAAGCAAGCTA 257
 226 GluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGluGlyValGln 245
 258 GAACTCTTTCGAGAGACTTCAGGTCACATTAAGAGCTTAAGAGGAGAAATATGCAG 317
 246 ArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeu 265
 318 CAGGCACGAGAGGAGAGAGACGCTAAAGAAAGCTCATCGAACAGTACCATTTCCAGATTACCTTA 377
 266 AsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGlu 285
 378 AATGAAATTCAGCCAGCTGGAGCGAGTACATCCCAAGCCAACTCCGACAGAA 437
 286 AsnMetGluLeuAlaGluLysLysLeuIleGluGlnTyrGluLeuArgGluGlu 305
 438 AACATTGAGCTGGGGGAGAGAGCTAAAGAAAGCTCATCGAACAGTACCGACTCAGGGAAGAG 497
 306 HisIleAspLysValPheLysHisLysAspLeuGlnGlnLeuValAspAlaLysLeu 325
 498 CACATTGATAAGGTGTTCAACATTAAGAACTGCAACAGCAGCTCGTGATGCCAACTG 557
 326 GlnGlnAlaGlnGluMetLysGluAlaGluGluArgHisGlnArgGluLysAspPhe 345
 558 CAGCAACGACACACAACTGATAAAGAGCTGATGATAAAGCATCAGAGAGAGAGAGATT 617
 346 LeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThr 365
 618 TTATTAAGAGAGAGCAGCAGAAATCGAGGCAACAAATACGAAACAAATGAAACAGCAAGATGA 677
 366 HisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeu 385
 678 CAACTTAACACAGCAGCTTCTCTTTATGATGATGATGATGATGATGATGATGATGATGATG 736
 386 SerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThr 403
 737 GCAAAAGCAATGAATGTTTACAACTTCAGACAGGAAATGGAANAGATGACN 790

CB247697 784 bp mRNA linear EST 09-JUL-2003
 UI-M-FIO-cea-m-20-0-UI-r1 NIH_BMAP_F10 Mus musculus cDNA clone
 IMAGE:6837021 5', mRNA sequence.
 CB247697
 ACCESSION CB247697.1 GI:28369341
 VERSION
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 784)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-rc@mail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov

This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pYX-5.

FEATURES

Location/Qualifiers

1..784

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:6837021"

/tissue_type="whole brain"

/dev_stage="embryo 12.5dpc"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NTH BMAP F10"

/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CAGCAGCAGC. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:

Pred. No.: 2,81e-54 Length: 784

Score: 972.00 Matches: 201

Percent Similarity: 87.66% Conservative: 5

Best Local Similarity: 85.53% Mismatches: 27

Query Match: 35.97% Indels: 3

DB: 14 Gaps: 2

US-10-023-523-8 (1-530) x CB247697 (1-784)

QY 1 LysSerProGlyGlnProGluAlaGlyProGluGlyValaGlnGluArgProSerGln 20

Db 83 AAGGGGAGCCGGGCGGACCGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 142

QY 21 AlaAlaProAlaValGluAlaGluGlyProGlySerGlnAlaProArgLysProGlu 40

Db 143 ACAGCTCTCGGGCGGGAAGCCGAAGGT---TCCACAGCCAGGCTCTGGGAAGCCGAG 199

QY 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 60

Db 200 GGGGCTCGAGTAAAGCAGCTCAGCTGGGGGCGCTCTGTACCTCTCTGAGGAGCTGAGC 259

QY 61 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnGlnGlyGlyProGly 80

Db 260 CGGCAGTTGGAAGACATCTGTAGTACATCTGTGTGACACACATCAGGAGGCGCGCT 319

QY 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGlyLysSerArgThrTyr 100

Db 320 GAGGAGGAGCAGAGGTCAGCCCTGAGCCGCGAAGACACCGAGAGTCCCGAACCTAT 379

QY 101 ValAlaArgAsnGlyGluProGluPro---ThrProValValTyrGlyGlyLysGluPro 119

Db 380 GCAGCCAGGAGTGGGAGCTGACCCAGGCATTCAGTCTGTCACCGCGAGAGGAGACC 439

QY 120 SerLysGlyAspProAlaThrGluGluIleArgGlnSerAspGluValGlyAspArgAsp 139

Db 440 TCTAAGGAGGAGCTGGAAACAGAGGAGATCCGAGCGAGTGTAGTTCGAGACCCGAGAC 499

QY 140 HisArgArgProGlnGluLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeu 159

Db 500 CATCGGAGGCCACAGAGAGAGAGAGAGCCAGGGTCTTAGGAGGAGAGATCACTCTGCTG 559

QY 160 MetGlnThrIleuAsnThrIleuSerThrProGluGluLysLeuAlaAlaLeuCysLysLys 179

Db 560 ATGCAGACACTGACACGCTGAGTATCCCCAGAGGAGAGCTGCTGCACTGTGCAAGAAG 619

QY 180 TyrAlaGluLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLys 199

Db 620 TATGCTGAGCTGCTGGAGAGCATCGGAATCTCGAGAGCAGATGAGCTCTCTGCAGAG 679

QY 200 LysGlnSerGlnLeuValGlnGlnLysAspHisLeuArgGlyGluHisSerLysAlaVal 219

Db 680 AAGCAGAGCAGCTGCTGTCAGGAGAGGA-CATCTCGGANGGAAACACAGCAGGCTGTC 738

QY 220 LeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArg 234

Db 739 CTGGCCCGAAGCAGCTGAGAGTCTGTGCCCGAGGCTGCACCG 783

RESULT 50

BG975353

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

796 bp mRNA linear EST 12-JUN-2001

602842995F1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:4978775 5', mRNA sequence.

BG975353

BG975353.1 GI:14362990

EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 796)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LLAM10975 row: n column: 24

High quality sequence stop: 780.

Location/Qualifiers

1..796

/organism="Mus musculus"

/mol_type="mRNA"

/strain="NMRI"

/db_xref="taxon:10090"

/clone="IMAGE:4978775"

/tissue_type="tumor, gross tissue"

/dev_stage="5 months"

/lab_host="DH10B"

/clone_lib="NCI_CGAP Mam4"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Priscilla Furth, NIH Reference for transgenic model: Li et al., Cell Growth and Differentiation 7, 3-11 (1996)."

FEATURES

source

ORIGIN

Alignment Scores:

Pred. No.: 5,65e-54 Length: 796

Score: 967.50 Matches: 199

Percent Similarity: 86.55% Conservative: 7

Best Local Similarity: 83.61% Mismatches: 28

Query Match: 35.81% Indels: 4

DB: 12 Gaps: 3

US-10-023-523-8 (1-530) x BG975353 (1-796)


```
Qy 1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 20
Db |||||
78 AAGGCAGCCCGGGGCAACCGGAGAGCAGGACCGGAGGAGCCCATGGACACCCAGACAG 137
Qy 21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40
Db |||||
138 ACAGCTCTCTGGGGCGGAGCGGAAGGT--TCCACAAGCCAGGCTCTCTGGGAAGACCGAG 194
Qy 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 60
Db |||||
195 GGGGCTCGAGCTAAAGCAGCTCAGCCTGGGGCCCTCTGTGACGTCTCTGAGGAGCTGAGC 254
Qy 61 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyProGly 80
Db |||||
255 CGGCAAGTTGGAGACATCTCTGATACATCTGTGTGGACAACATCAGGAGGCGCCGCT 314
Qy 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
Db |||||
315 GAGGAGGAGCAGCAGGCTGAGCCACTGAGCCGGAAGACACGAGAAAGTCCCGAACCTAT 374
Qy 101 ValAlaArgAsnGlyGluProGluPro---ThrProValValTyrGlyGluLysGluPro 119
Db |||||
375 GCAGCCAGCATGGGGAGCCCTGAACCCAGGATTCAGTCTGTCACGCGGAGAGGAGACC 434
Qy 120 SerLysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAsp 139
Db |||||
435 TCTAAGGAGAGCCTGGAAACAGAGAGATCCGAGCAAGTGATGAGTTGGAGACCGAGAC 494
Qy 140 HisArgArgProGlnGluLysIleLysAlaLysGlyLeuGlyLysGluIleThrLeuLeu 159
Db |||||
495 CATCGAGGCCACAGGAGAGAAAGCAAGGCTTAGGGAAGGAGATCACTCTGCTG 554
Qy 160 MetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaAlaLeuCysLysLys 179
Db |||||
555 ATGCAAGACTGACACACCTGATACCCAGAGGAGAGCTGGCTGCACTGTGCAAGAAG 614
Qy 180 TyrAlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeu---LeuGln 198
Db |||||
615 TATGCTGAGCTGCTGGAAGAGCATCGAACTCGCAGAAGCAGCATGAAGCTCCTTGCAGA 674
Qy 199 LysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGlu-HisSerLysAl 218
Db |||||
675 ACGAAGCAGAGCCAGCTCTGTGAGGAGAGAGGACCATCTGCGAGGGGAACCAAGCAGCAGGC 734
Qy 218 aValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHis 235
Db |||||
735 TGTCTGCCCCGAGCAAGCTTGAGAGTCTGTGCGCGGAGCTGCAAGGGCAC 786
```

Search completed: June 8, 2004, 12:35:08
Job time : 3911.61 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 8, 2004, 06:37:54 ; Search time 598.959 Seconds
(without alignments)
3759.096 Million cell updates/sec

Title: US-10-023-523-8

Perfect score: 2702

Sequence: 1 KSSPQBPAGPEGAQERSQ.....APSTEASGQTGPQPTARS 530

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 50 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US10023523/runat 08062004 063748 29931/app query.fasta_1.1422
-DB=N Geneseq 29Jan04 -OFT=fastac -SUFFIX=ring -MINMATCH=0.1 -LOOPECL=0
-LOOPEXT=0 -UNITS=bits -SPART=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=50 -DOALIGN=200 -THR SCORE=rcr -THR MAX=100 -THR MIN=0 -ALIGN=50
-MODE=LOCAL -OUTFMT=pcr -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10023523@cgn 1.1.755@runat 08062004 063748 29931 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPEXT=10 -XGAPEXT=0.5 -FGAPOPEXT=6
-FGAPEXT=7 -YGAPOPEXT=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq 29Jan04: *
1: Geneseqn1986s: *
2: Geneseqn1990s: *
3: Geneseqn2000s: *
4: Geneseqn2001as: *
5: Geneseqn2001bs: *
6: Geneseqn2002as: *
7: Geneseqn2003as: *
8: Geneseqn2003bs: *
9: Geneseqn2003cs: *
10: Geneseqn2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2702	100.0	4697	2 AAV32839	AAV32839 Human low
2	2702	100.0	4697	5 AAH26496	AAH26496 Human low
3	2693	99.7	1638	5 AAH26501	AAH26501 Human low
4	2693	99.7	2523	4 ABA09008	ABA09008 Human LDL
5	2693	99.7	2523	4 ABA09008	ABA09008 Human LDL
6	2693	99.7	5085	7 ACC46153	ACC46153 Human dit
7	2693	99.7	6577	7 ACC46130	ACC46130 Human dit
8	2465.5	91.2	4722	2 AAV32836	AAV32836 Rabbit lo

9	2465.5	91.2	4722	5	AAH26491	Rabbit lo
10	2343.5	86.7	5284	5	AAH26491	Rabbit lo
11	1754.5	64.9	2188	4	AAK51890	Human pol
12	1727	63.9	1793	7	ABZ34807	Coding se
13	1688	62.5	1793	3	AAH35218	Human ade
14	1688	62.5	1793	3	AAH35218	Human ade
15	1688	62.5	1793	3	AAH35218	Human ade
16	1688	62.5	1793	3	AAH35218	Human low
17	1688	62.5	1793	3	AAH35218	Human low
18	1688	62.5	1793	3	AAH35218	Human low
19	1688	62.5	1793	3	AAH35218	Human low
20	1688	62.5	1793	3	AAH35218	Human low
21	1688	62.5	1793	3	AAH35218	Human low
22	1688	62.5	1793	3	AAH35218	Human low
23	1688	62.5	1793	3	AAH35218	Human low
24	1688	62.5	1793	3	AAH35218	Human low
25	1688	62.5	1793	3	AAH35218	Human low
26	1688	62.5	1793	3	AAH35218	Human low
27	1688	62.5	1793	3	AAH35218	Human low
28	1688	62.5	1793	3	AAH35218	Human low
29	1688	62.5	1793	3	AAH35218	Human low
30	1688	62.5	1793	3	AAH35218	Human low
31	1688	62.5	1793	3	AAH35218	Human low
32	1688	62.5	1793	3	AAH35218	Human low
33	1688	62.5	1793	3	AAH35218	Human low
34	1688	62.5	1793	3	AAH35218	Human low
35	1688	62.5	1793	3	AAH35218	Human low
36	1688	62.5	1793	3	AAH35218	Human low
37	1688	62.5	1793	3	AAH35218	Human low
38	1688	62.5	1793	3	AAH35218	Human low
39	1688	62.5	1793	3	AAH35218	Human low
40	1688	62.5	1793	3	AAH35218	Human low
41	1688	62.5	1793	3	AAH35218	Human low
42	1688	62.5	1793	3	AAH35218	Human low
43	1688	62.5	1793	3	AAH35218	Human low
44	1688	62.5	1793	3	AAH35218	Human low
45	1688	62.5	1793	3	AAH35218	Human low
46	1688	62.5	1793	3	AAH35218	Human low
47	1688	62.5	1793	3	AAH35218	Human low
48	1688	62.5	1793	3	AAH35218	Human low
49	1688	62.5	1793	3	AAH35218	Human low
50	1688	62.5	1793	3	AAH35218	Human low

ALIGNMENTS

RESULT 1
AAV32839
ID AAV32839 standard; cDNA; 4697 BP.

XX AC AAV32839;

XX DT 09-NOV-1998 (first entry)

XX DE Human low density lipoprotein binding protein LBP-3 cDNA.

XX KW Low density lipoprotein binding protein; LDL binding protein 3; LBP-3;
XX KW receptor; human; atherosclerosis; diagnosis; therapy; vaccine; ss.

XX OS Homo sapiens.

XX PH Key Location/Qualifiers

XX FT CDS 3..1595

XX FT misc_feature /tag= a

XX FT misc_feature /tag= b

XX FT misc_feature /note= "Claim 23"

XX DN WO9823282-A1.

XX XX 04-JUN-1998.

XX PF 26-NOV-1997; 97WO-US021857.

XX 27-NOV-1996; 96US-0031930P.
 PR 03-JUN-1997; 97US-0048547P.
 XX
 XX (BOST-) BOSTON HEART FOUND INC.
 PI Lees AM, Lees RS, Law SW, Arjona AA;
 XX WPI; 1998-322455/28.
 DR P-PSDB; AAW49042.
 XX
 PT Nucleic acid encoding low density lipoprotein binding proteins and
 PT related vectors - transformed cells, proteins, and modulators of binding,
 PT useful for treatment and diagnosis of atherosclerosis and for identifying
 PT subjects at risk.
 XX
 PS Claim 10; Fig 17; 47pp; English.
 CC
 CC This cDNA clone codes for novel human low density lipoprotein (LDL)
 CC binding protein LBP-3 (see AAW49042). It was isolated by screening human
 CC liver, aorta and foetal brain cDNA libraries with rabbit LBP-3 cDNA. cDNA
 CC clones (see AAW32834-39) and encoded rabbit and human LBPs (see AAW49037-
 CC 42) are claimed. An abnormality in LBP metabolism or structure is
 CC diagnostic of a risk for atherosclerosis. The invention provides: methods
 CC for determining if an animal is at risk for atherosclerosis (e.g. for
 CC prenatal screening); methods for treating atherosclerosis (including gene
 CC therapy) using e.g. LBP polypeptides to bind LDL and thereby prevent
 CC formation of atherosclerotic plaque; and methods for treating a cell
 CC having an abnormality in LBP structure or metabolism. Pharmaceutical and
 CC vaccine compositions are also provided, as well as recombinant vectors
 CC and host cells used to produce recombinant LBP
 XX
 SQ Sequence 4697 BP; 1097 A; 1198 C; 1289 G; 1113 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1.4e-157 Length: 4697
 Score: 2702.00 Matches: 530
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0

US-10-023-523-8 (1-530) x AAW32839 (1-4697)

QY 1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 20
 DB 3 AAWAGCAGCCAGGACAAACCGAAGCAGGAGCCGAGGAGCCAGGAGCCGAGCCAGCCAG 62
 QY 21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40
 DB 63 GCGCCTCCTGCAGTAGAAGCAGAGAGTCCCGGAGCAGCAGCCAGCTCTCGAAGCCGGAG 122
 QY 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaAlaArgAspValSerGluGluLeuSer 60
 DB 123 GGGGCTCAAGCCAGACAGCGGTCTAGTCTGGGGCCCTTCGTGATGCTCTGAGGAGCTGAGC 182
 QY 61 ArgGlnLeuGluAspIleLeuSerThrTyCysValAspAsnGlnGlyGlyProGly 80
 DB 183 CGCCAACTGGAGACATACCTAGCACAATCTGTGTGACATTAACCCAGGGGGGCCCGGAC 242
 QY 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluSerArgThrTy 100
 DB 243 GAGGATGGGCGCAGGAGTGGAGCGGTGAGCCCGGTGAAACCGAAGATGCGAGAAGTCCCGGACCTAT 302
 QY 101 ValAlaArgAsnGlyGluProGluProThrProValValTyGlyGluLysGluProSer 120
 DB 303 GTGCGAAGGATGGGAGGCTGAACTCACTCCAGTAGTCTTAGGAGAGGAAGACCTCC 362
 QY 121 LysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHis 140
 DB 363 AAGGGGGATCCAAACACAGAGAAGATCCGCGCAGGTGAGGAGTCCGAGACCCGAGACCAT 422
 QY 141 ArgArgProGlnGluLysLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMet 160

DB 423 CGAAGGCCACAGAGAGAAAAAAGCCAGGGTTTGGGAAGGAGATCATGTTGCTGATG 482
 QY 161 GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaAlaLeuCysLysLysTy 180
 DB 483 CAGACATTGAATACTCTGAGTACCCAGAGGAGAACTGGCTCTCTGTGCAAGAGTAT 542
 QY 181 AlaGluLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuGlnLysLys 200
 DB 543 GCTGAACCTGCTGAGGAGCACCAGGAATTCACAGACGATGAGCTGTCTACAGAAAAA 602
 QY 201 GlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGlyGlnHisSerLysAlaValLeu 220
 DB 603 CAGAGCCAGCTGCTGCAAGAGAGGACCACTCGCGGTGAGCAGACAGCAAGCCGCTCTG 662
 QY 221 AlaArgSerLysLeuGlnSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLys 240
 DB 663 GCCCGCAGCAAGCTTGAGAGCCTATGCGCTGAGCTGACGCGGCACACCGCTCCCTCAAG 722
 QY 241 GluGluGlyValGlnArgAlaArgGluGluGluGluLysArgLysGluValThrSerHis 260
 DB 723 GAAGAAGGTGTCAGCGGCGCCCGGAGGAGGAGGAGAACCGCAAGAGGTGACCTCGCAC 782
 QY 261 PheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSer 280
 DB 783 TTCCAGGTGACACTGAATGACATTCAGCTGCAGATGGAACAGCACATGAGCGCAACTCC 842
 QY 281 LysLeuArgGlnGluAsnMetGluLeuAlaGluArgLysLysLysLeuIleGluGlnTy 300
 DB 843 AAGCTGCGCCAAAGAACATGAGCTGCTGAGAGGCTCAAGAGCTGATGAGCAGTAT 902
 QY 301 GluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeu 320
 DB 903 GAGCTGGCGAGGACATATCGACAAAGTCTCAACACAGGACCTACACAGCAGCTG 962
 QY 321 ValAspAlaLysLeuGlnGlnAlaGlnMetLeuLysGluAlaGluGluArgHisGln 340
 DB 963 GTGGATGCCAAGCTCCAGCAGCCCGAGGAGATGCTTAAAGGAGGAGAGAGCGGCCAG 1022
 QY 341 ArgGluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMet 360
 DB 1023 CGGAGAGAGGATTTCTCTCTGAAAGAGCAGTAGTCTCCAGAGAGATGTGAGCTGATG 1082
 QY 361 LysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyThrGluLysPheGluGlu 380
 DB 1083 AAGCAGCAGAGAGACCCACCTGAAGCAACAGCTTGCCTTATACACAGAGAGATTGAGCAG 1142
 QY 381 PheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGlu 400
 DB 1143 TTCCAGAACACACTTTCGAAAGCAGCGAGGTATTCACCATTTCAAGCAGAGATGGAA 1202
 QY 401 LysMetThrLysLysLysLysLeuGluLysGluThrThrMetTyArgSerArgTrp 420
 DB 1203 AAGATGACTAAGAAGATCAAGAAGCTGGAGAAAGAACACCATGTACCGGTCCCGGTGG 1262
 QY 421 GluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGlu 440
 DB 1263 GAGAGCAGCAACAGGCGCTCTTGAGATGCTGAGGAGAAACAGTCCGGGATAAAGAA 1322
 QY 441 LeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThr 460
 DB 1323 CTCGAGGCGCTGCGAGGTAAATAATCCACGCTGGAGAGAGCTGTGCGGGCAGTCGACACA 1382
 QY 461 GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlnGlySerLeu 480
 DB 1383 GAGCGCAATGACCTGAACAGAGGTACAGGACTGAGTCTGCTGGTGGCCAGGCTCCCTC 1442
 QY 481 ThrAspSerGlyProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerPro 500
 DB 1443 ACTGACAGTGGCCCTGAGAGGAGGCGGAGGCGCTGGGCTCAAGCACCAGCCTCCCC 1502
 QY 501 ArgValThrGluAlaProCysTyThrProGlyAlaProSerThrGluAlaSerGlyGlnThr 520

DB 1503 AGGGTCACAGAGCGCTTGTCTACCCAGGAGCACCAGCAGACAGATCATGGCCAGACT 1562
 QY 521 GlyProGlnGluProThrSerAlaArgAla 530
 DB 1563 GGGCTCAGAGCCACCTCCCGAGGGCC 1592

RESULT 2

AAH26496
 ID AAH26496 standard; cDNA; 4697 BP.

AC AAH26496;

DT 12-NOV-2001 (first entry)

DE Human low density lipoprotein binding protein 3 (LBP-3) cDNA.

XX Low density lipoprotein binding protein 3; LBP-3; LDL; human;
 KW atherosclerosis; antiarteriosclerotic; gene therapy; diagnosis; vaccine;
 XX ss;

OS Homo sapiens.

FH Key Location/Qualifiers
 FT CDS 3..1596
 FT /*tag= a
 FT /partial

XX WO200164874-A2.

PN 07-SEP-2001.

XX 28-FEB-2001; 2001WO-US006356.

XX 02-MAR-2000; 2000US-00517849.

PR 14-JUL-2000; 2000US-00616289.

XX (BOST-) BOSTON HEART FOUND INC.

XX Lees AM, Lees RS, Law SW, Arjona AA;

XX WPI; 2001-565505/63.

DR P-PDB; AAB82804.

XX New isolated low density lipoprotein binding polypeptide for treating,
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.

XX Example 4; Fig 17; 143pp; English.

XX The present sequence is that of a partial cDNA encoding novel human low
 CC density lipoprotein binding protein 3 (LBP-3, see AAB82804). Clones were
 CC isolated from human foetal brain, liver and aorta cDNA libraries using
 CC rabbit LBP-3 cDNA as probe. A full-length cDNA clone is given in
 CC AAH26501, and a genomic DNA sequence is given in AAH26497. LBP-3 nucleic
 CC acids are among claimed polynucleotides of the invention that encode
 CC novel polypeptides capable of binding to native and methylated LDL. Also
 CC claimed are isolated LBP polypeptides, and biologically active fragments
 CC and analogues of them, as well as expression vectors, cells and methods
 CC of producing the LBPs. Methods of determining if an animal is at risk for
 CC atherosclerosis, methods for evaluating an agent for use in treating
 CC atherosclerosis, and methods for treating a cell having an abnormality in
 CC structure or metabolism of LBP are claimed. Pharmaceutical compositions
 CC comprising an LBP polypeptide or nucleic acid, and vaccine compositions,
 CC are also claimed

SQ Sequence 4697 BP; 1097 A; 1197 C; 1290 G; 1113 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.4e-157 Length: 4697
 Score: 2702.00 Matches: 530
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 5

DB:

US-10-023-523-8 (1-530) x AAH26496 (1-4697)
 QY 1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 20
 DB 3 AAANGCAGCCAGGACAAACCGGAGCAGGACCCGAGGAGCCAGGAGCCGAGCCAGCAG 62
 QY 21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40
 DB 63 GCGGCTCTCGTAGAAGCAGAGAGGTCCCGGAGCAGCAGCCAGGCTCTCGGAAGCCGAG 122
 QY 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaAlaLeuArgAspValSerGluLeuSer 60
 DB 123 GGGCTCAAGCCAGACCGCTCAGCTCGGGCCCTTCGTGATGCTCTGAGGAGCTGAGC 182
 QY 61 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGly 80
 DB 183 CGCCAACTGGGAAGACATACCTAGCAGCACATACGTGTGGACAATAACACAGGGGGCCCCGGC 242
 QY 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
 DB 243 GAGGATGGGGCACAGGGTGAGCGGGCTGAACCCGAGATCGAGAGAAAGTCCCCGACCTAT 302
 QY 101 ValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGluLysGluProSer 120
 DB 303 GTGGCAAGGAATGGGAGCCCTGACCACTCCAGTAGTCTATCGAGAGAGGAACCTCC 362
 QY 121 LysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHis 140
 DB 363 AAGGGGATCCAAACACAGAGAGATCCGGCAGATGACGAGGTCCGAGACCGAGACCAT 422
 QY 141 ArgArgProGlnGluLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMet 160
 DB 423 CGAAGGCCACAGGAGAGAAAGAACCCAGGGTTTGGGAGAGGATCAGTGTCTGATG 482
 QY 161 GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuLysLysLysTyr 180
 DB 483 CAGACATTGAATACTCTGAGTACCCAGAGAGAGAGCTGCTGCTGTGCAAGAGTAT 542
 QY 181 AlaGluLeuLeuGluHisArgAsnSerGlnLysGlnMetLysLeuGlnLysLys 200
 DB 543 GCTGAATCTGTGGAGGAGCACCGGAATTCACAGAGCAGATGAAGTCTCTACAGAAAAG 602
 QY 201 GlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeu 220
 DB 603 CAGAGCCAGCTGTGTGCAAGAGAGACCACTGCGCGTGCAGCACACCAAGGCGTCTG 662
 QY 221 AlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLys 240
 DB 663 GCCCGCAGCAGCTTGAGAGCCTATGCCGTGAGCTGCAGCGGCACCAACCCCTCCCTCAAG 722
 QY 241 GluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHis 260
 DB 723 GAAGAAGGTGTGACGCGGCGCCGAGAGAGAGAGAGAGCGCAAGGAGGTGACCTCGCAC 782
 QY 261 PheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSer 280
 DB 783 TTCAGGTGACACTGAATGACATTCAGTTCAGTGCAGATGGAACAGCACAATGAGCGCAACTCC 842
 QY 281 LysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTyr 300
 DB 843 AAGTCGCGCCAGAGAACATGGAGCTGGCTGAGAGGCTCAAGAGCTGATTGAGCAGTAT 902
 QY 301 GluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeu 320
 DB 903 GAGCTGCGGAGGAGCATATCGACAAAGTCTTCAACACAGAGGACCTTACACAGCAGCTG 962
 QY 321 ValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGln 340
 DB 963 GTGATGCCAAGCTCCAGAGGCCCGAGAGATGCTAAAGAGGGGAGAGGAGCGGACACAG 1022
 QY 341 ArgGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMet 360

1023 CGGAGAGAGATTTCTCTCTGAAAGAGGAGTAGAGTCCAGAGAGTGTGTGAGTGATG 1082
 QY LysGlnGlnGlnThrHisLeuLysGlnGlnLeuAlaLeuThrThrGluLysPheGluGlu 380
 1083 AAGCAGCAAGAGACCCACCTGAAGCAACAGCTGCTCCCTATACACAGAGAAGTTTGAGGAG 1142
 QY PheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGlu 400
 1143 TTCAGAACACATTTTCCAAAAGCAGCGAGGTATTCACACATTCAAGCAGAGATGGA 1202
 QY LysMetThrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 420
 1203 AAGATGACTAAGAGATCAAGAGCTGGAGAGAAAGAAACCAACATGTACCGGTCCCGGTG 1262
 QY GluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGlu 440
 1263 GAGAGCAGCAACAAAGCCCTGCTTGATGGCTGAGGAGAAACACAGTCCGGATTAAGAA 1322
 QY LeuGluGlyLeuGlnValLysLysLysLysLysLysLysLysLysLysLysLysLys 460
 1323 CTGGAGGGCTCGAGGATTAATCCACCGCTGGAGAGCTGTCCGGGCACTGCAGACA 1382
 QY GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeu 480
 1383 GAGCGCAATCACCTGAACAAGAGGGTACAGGACCTGAGTGTCTGGTGGCCAGGCTCCCTC 1442
 QY ThrAspSerGlyProGluArgArgProGluGlyProGlyValAlaGlnAlaProSerPro 500
 1443 ACTGACAGTGGCCCTGAGAGGAGGCCAGAGGGGCTTGGGGCTCAAGCACCCAGCTCCCC 1502
 QY ArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThr 520
 1503 AGGTCACAGAGCGCTTGTACCCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1562
 QY GlyProGlnGluProThrSerAlaArgAla 530
 1563 GGGCCTCAAGAGCCACCTCCCGCAGGGCC 1592
 RESULT 3
 AAH26501
 ID AAH26501 standard; DNA; 1638 BP.
 AC AAH26501;
 XX
 DT 12-NOV-2001 (first entry)
 DE Human low density lipoprotein binding protein 3 (LBP-3) DNA.
 DE
 XX Low density lipoprotein binding protein 3; LBP-3; LDL; human;
 KW atherosclerosis; antiarteriosclerotic; gene therapy; diagnosis; vaccine;
 KW ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200164874-A2.
 XX
 PD 07-SEP-2001.
 XX
 XX 28-FEB-2001; 2001WO-US006356.
 XX
 PR 02-MAR-2000; 2000US-00517849.
 PR 14-JUL-2000; 2000US-00616289.
 XX
 PA (BOST-) BOSTON HEART FOUND INC.
 XX
 XX Lees AM, Lees RS, Law SW, Arjona AA;
 XX
 XX WPI; 2001-565505/63.
 DR P-PSDB; AAB2808.
 DR
 XX New isolated low density lipoprotein binding polypeptide for treating,
 PT diagnosing and/or identifying therapeutic agents for atherosclerosis.
 XX

PS Claim 8; Fig 8A; 143pp; English.
 XX
 CC The present sequence is that of the coding region of the human gene (see
 CC also AAH26496) encoding novel human low density lipoprotein binding
 CC protein 3 (LBP-3, see AAB2808). The gene was isolated from a genomic DNA
 CC library using LBP-3 cDNA as probe. The LBP-3 protein predicted from the
 CC present sequence differs from that predicted from a cDNA clone (see
 CC AAB2804) in that it contains an additional 16 amino acids at its N-
 CC terminus (the cDNA is a 5' truncation) and has Asn at amino acid position
 CC 130 rather than Tyr. LBP-3 nucleic acids are among claimed
 CC polynucleotides of the invention that encode novel polypeptides capable
 CC of binding to native and methylated LDL. Also claimed are isolated LBP
 CC polypeptides, and biologically active fragments and analogues of them, as
 CC well as expression vectors, cells and methods of producing the LBPs.
 CC Methods of determining if an animal is at risk for atherosclerosis, and
 CC methods for evaluating an agent for use in treating atherosclerosis, and
 CC methods for treating a cell having an abnormality in structure or
 CC metabolism of LBP are claimed. Pharmaceutical compositions comprising an
 CC LBP polypeptide or nucleic acid, and vaccine compositions, are also
 CC claimed
 XX
 SQ Sequence 1638 BP; 490 A; 417 C; 520 G; 211 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1.59e-157 Length: 1638
 Score: 2693.00 Matches: 529
 Percent Similarity: 99.81% Conservative: 0
 Best Local Similarity: 99.81% Mismatches: 1
 Query Match: 99.67% Indels: 0
 DB: 5 Gaps: 0
 US-10-023-523-8 (1-530) x AAH26501 (1-1638)
 QY 1 LysSerSerProGlyGlnProGluAlaGlyProGluValAlaGlnGluArgProSerGln 20
 DB 49 AAAGCAGCCAGGACCAACCGAAGCAGGACCCAGGGAGCCAGAGGGGCCACCCAG 108
 QY 21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40
 DB 109 GCGGCTCTCGTAGAAGCAGAGAGTCCCGCAGCAGCGAGCTCTCTCGAAGCGCGAG 168
 QY 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 60
 DB 169 GGTGCTCAAGCCAGACCGGTCTGAGTCTGGGGCCCTTCGTGATGCTCTCTGAGGAGCTGAGC 228
 QY 61 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGly 80
 DB 229 CGCCAACTGGAGACATACCTAGCAGCATACTGTGTGACAAATAACCCAGGGGGCCCCGGC 288
 QY 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
 DB 289 GAGGATGGGCACAGGGTGAGCCGGCTGAACCCGAAAGATCGAGAGAAGTCCCGGACCTAT 348
 QY 101 ValAlaArgAsnGlyGluProGluProThrProValTyrGlyGlyLysGluProSer 120
 DB 349 GTGCGAAGGATGGGAGCTGAAACCACTCCAGTAGTCAATCGAGAGAGAAACCTCC 408
 QY 121 LysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHis 140
 DB 409 AAGGGGGATCAAAACACAGAAAGAGATCCCGCAGAGTACAGAGTCCGAGAGCCAGACCAT 468
 QY 141 ArgArgProGlnGlyLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMet 160
 DB 469 CGAAGGCCACAGGAGAGAGAAAAGCCAGGGTTGGGTAAAGAGATCAGCTTGTCTGATG 528
 QY 161 GlnThrLeuAsnThrLeuSerThrProGluGlyLysLeuAlaLeuCysLysLysTyr 180
 DB 529 CAGACATTGAATATCTCTGAGTACCCAGAGAGAGCTGGCTCTCTCTGTCAAGAGATAT 588
 QY 181 AlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLys 200
 DB 589 GCTGACTGCTGGAGGAGCAGCCGGAATTCAGAGCAGATGAGGCTCTCTACAGAAAAG 648

QY 461 GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeu 480
 DB 1547 GAGGCAATGACCTGAAACAGAGGGGTACAGACCTGAGTGTGGTCCAGGGCTCCCTC 1606
 QY 481 ThrAspSerGlyProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerPro 500
 DB 1607 ACTGACAGTGGCCCTGAGAGAGGCGCCAGAGGGGCTGGGGCTCAAGCACCAGCTCCCTC 1666
 QY 501 ArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThr 520
 DB 1667 AGGTCACAGAGCGCTTGTCTCCAGGAGCACCAGAGCAGACAGAGCATCAGGCCAGACT 1726
 QY 521 GlyProGlnGluProThrSerAlaArgAla 530
 DB 1727 GGGCTCAAGAGCCACCTCCGCGAGGGCC 1756
 RESULT 6
 ID ACC46153 standard; cDNA; 5085 BP.
 XX
 AC ACC46153;
 XX
 DT 02-JUN-2003 (first entry)
 XX
 DE Human dithp extracellular signalling protein-encoding cDNA.
 XX
 KW Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;
 KW cancer; cell proliferative disorder; autoimmune disorder;
 KW inflammatory disorder; infection; hormonal disorder; metabolic disorder;
 KW neurological disorder; gastrointestinal disorder; transport disorder;
 KW connective tissue disorder; drug screening; proteome analysis;
 KW gene therapy; antisense therapy; genotyping; transgenic animal; knock in;
 KW disease model; toxicological testing; transcript imaging;
 KW extracellular signalling; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FN WO200297031-A2.
 XX
 PD 05-DEC-2002.
 XX
 PF 27-MAR-2002; 2002WO-US010056.
 XX
 PR 28-MAR-2001; 2001US-0279619P.
 PR 29-MAR-2001; 2001US-0280067P.
 PR 29-MAR-2001; 2001US-0280068P.
 PR 16-MAY-2001; 2001US-0291280P.
 PR 17-MAY-2001; 2001US-0291829P.
 PR 17-MAY-2001; 2001US-0291849P.
 PR 19-JUN-2001; 2001US-0299428P.
 PR 20-JUN-2001; 2001US-0299776P.
 PR 20-JUN-2001; 2001US-0300001P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
 PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Anshey SR;
 PI Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;
 PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
 PI Flores V, Marwaha R, Lo A, Lan RV, Urashka ME;
 XX
 DR WPI; 2003-129518/12.
 DR P-PSDB; ABR41210.
 XX
 PT Novel human diagnostic and therapeutic polypeptide useful for identifying
 PT test compound which specifically binds to a polypeptide encoded by human
 PT diagnostic and therapeutic polynucleotide, and to induce antibodies.
 XX
 PS Claim 2; SEQ ID NO 74; 591pp; English.
 XX
 CC The invention relates to novel human diagnostic and therapeutic
 CC polynucleotides designated dithp (ACC46080-ACC46749) and to their encoded
 CC proteins (DITHP; ABR41136-ABR41812). The invention also relates to

CC polynucleotide sequences at least 90% identical to the dithp cDNA
 CC sequences of the invention; recombinant vectors, host cells and
 CC transgenic organisms comprising a dithp nucleic acid sequence; the
 CC recombinant production of DITHP proteins; antibodies specific for DITHP
 CC proteins; microarrays comprising dithp nucleic acid sequences; methods of
 CC detecting dithp nucleotide and protein sequences; methods of screening
 CC for compounds which specifically bind a DITHP protein; and methods of
 CC assessing the toxicity of test compounds using a dithp hybridisation
 CC probe. Dithp nucleic acid sequences and DITHP proteins may be used in the
 CC diagnosis of a wide variety of conditions including cancer and other cell
 CC proliferative disorders; autoimmune or inflammatory disorders; bacterial,
 CC viral, fungal or parasitic infections; hormonal disorders; metabolic
 CC disorders; neurological disorders; gastrointestinal disorders; transport
 CC disorders; and connective tissue disorders. They may also be used to
 CC screen for modulators of protein activity or gene expression. DITHP
 CC proteins can additionally be used in analysis of the proteome of a tissue
 CC or cell type and to induce antibodies. The dithp nucleic acids are
 CC additionally useful in somatic or germline gene therapy of the disorders
 CC mentioned above, as a source of antisense sequences, as a source of
 CC probes and primers, in genotyping and identification of individuals, in
 CC the generation of transgenic animal models of human disease or knock in
 CC humanised animals, in toxicological testing, and in transcript imaging.
 CC The present sequence represents a dithp cDNA encoding a DITHP protein
 CC which has extracellular signalling activity. Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 5085 BP; 1178 A; 1310 C; 1407 G; 1190 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 5, 5e-157 Length: 5085
 Score: 2693.00 Matches: 529
 Percent Similarity: 99.81% Conservative: 0
 Best Local Similarity: 99.81% Mismatches: 1
 Query Match: 99.67% Indels: 0
 DB: 7 Gaps: 0
 US-10-023-523-8 (1-530) x ACC46153 (1-5085)

QY 1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 20
 DB 328 AAAAGCAGCCCGAGACACCGGAAAGCAGGAGCCCGAGGAGCCCGAGCGCCAGCCAG 387
 QY 21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40
 DB 388 GCGGCTCTCGAGTAGAAGCAGAGAGGTCCCGGAGCAGCCAGGCTCTCTGGAAGCCGAG 447
 QY 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluLeuSer 60
 DB 448 GGGGCTCAAGCAGACAGCGCTCAGTCTGGGCGCCCTCGTGATGTCTCTGAGGAGCTGAGC 507
 QY 61 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnGlnGlyGlyProGly 80
 DB 508 GCCAACTGGGAAGACATCTAGCAGCATCTGTGGCACAATAACCGAGGGGGGGCCCGGC 567
 QY 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
 DB 568 GAGGATGGGACAGGGTGTAGCGCGCTGAACCCCAAGATGCAGAGAAAGTCCCGGACCTAT 627
 QY 101 ValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGluLysGluProSer 120
 DB 628 GTGGCAAGAAATGGGAGCCCTGAACCACTCCAGTAGTCAATGGAGAGAGGAAACCTCC 687
 QY 121 LysGlyAspProAsnThrGluIleArgGlnSerAspGluValGlyAspArgAspHis 140
 DB 688 AAGGGGATCAACACACAGAGAGATCCGGCAGAGTCAACGAGTCCGAGACCGAGACCAT 747
 QY 141 ArgArgProGlnGluLysLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMet 160
 DB 748 CGAAGGCCACAGGAGAGAAAAAGCCAGGGTTTGGGGAAGGAGATCACGTTGCTGATG 807
 QY 161 GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysTyr 180

Db 808 CAGACATTGAATCTCTGAGTACCCAGAGAGAGTGGCTCTGTGTCAAGAAGTAT 867
Qy |||||
Db 181 AlaGluLeuGluGluHsAsnSerGlnLysGlnMetLysLeuGluGlnLysLys 200
Qy |||||
Db 868 GCTCACTGCTGGAGGAGCAGCGAATTCAGAGAGCAGATGAAGCTCTACAGAAAG 927
Qy |||||
Db 201 GlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeu 220
Qy |||||
Db 928 CAGAGCCAGCTGGTGCAGAGAGAGACACCTGCGCGTGAGCAGCAGCAGAGCGCTCTG 987
Qy |||||
Db 221 AlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLys 240
Qy |||||
Db 988 GCCGCGAGCAGCTTGAGAGCCTATGCGTGAGCTGAGCGGCAGACACCTCCCTCAAG 1047
Qy |||||
Db 241 GluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHis 260
Qy |||||
Db 1048 GAAGAGGTGTGACGGCGGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1107
Qy |||||
Db 261 PheGlnValThrLeuAsnAspLysGlnLeuGlnMetGluGlnHisAsnGluArgAsnSer 280
Qy |||||
Db 1108 TTCAGGTGACCTGATGATGACATTCAGCTGCAGATGGAACAGCACATGAGCGCACTCC 1167
Qy |||||
Db 281 LysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuLysGlnTyr 300
Qy |||||
Db 1168 AAGCTGCGCCAGAGACATGAGCTGGCTGAGAGGCTCAAGAGCTGATTGAGCAGTAT 1227
Qy |||||
Db 301 GluLeuArgGluGluHisLysAspLysValPheLysHisLysLysLeuGlnGlnLeu 320
Qy |||||
Db 1228 GAGCTGCGCGAGAGCATATCGCAAGAGCTTCAAAACAGAGGACCTCAACAGCAGCTG 1287
Qy |||||
Db 321 ValAspAlaLysLeuGlnGlnAlaGlnGluMetLysLysGluAlaGluGluArgHisGln 340
Qy |||||
Db 1288 GTGGATGCCAAGCTCCAGAGGCGCCAGAGATGCTTAAGAGGAGCAGAGAGCGGACAG 1347
Qy |||||
Db 341 ArgGluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMet 360
Qy |||||
Db 1348 CCGGAGAGAGATTTCTCTGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1407
Qy |||||
Db 361 LysGlnGlnGlnThrHisLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGlu 380
Qy |||||
Db 1408 AACAGCAGAGAGACCCACCTGAGAGCAGCAGCTTGCCTTATACACAGAGAGTTGAGGAG 1467
Qy |||||
Db 381 PheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGlu 400
Qy |||||
Db 1468 TTCAGACACACATTTTCCAAAGCAGGAGGATTCACACATTCACAGCAGAGATGGAA 1527
Qy |||||
Db 401 LysMetThrLysLysLysLysLysLeuGluLysGluThrThrMetTyrArgSerArgTrp 420
Qy |||||
Db 1528 AAGATGACTAAGAAGATCAAGAGCTGGAAGAGAAACACCATGATGACCGGTCCCGGTGG 1587
Qy |||||
Db 421 GluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGlu 440
Qy |||||
Db 1588 GAGAGCAGCAACAGAGCCCTGCTTGAGATGGCTGAGAGAGAAACAGCTCCGGATTAAGNA 1647
Qy |||||
Db 441 LeuGluGluLeuValLysLysLysGlnArgLeuGluLysLeuCysArgAlaLeuGlnThr 460
Qy |||||
Db 1648 CTGGAGGGGCTGAGGATTAATAATCCACGCTGAGAGAGCTGCGGCACTGCAGACA 1707
Qy |||||
Db 461 GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaClyGlyGlnGlySerLeu 480
Qy |||||
Db 1708 GAGCGCAATGACTGAAACAGAGGGGTACAGGACCTGAGTGCTGGTGCGCAGGGCTCCCTTC 1767
Qy |||||
Db 481 ThrAspSerClyProGluArgArgProGluGlyProGlyValAlaGlnAlaProSerSerPro 500
Qy |||||
Db 1768 ACTGACAGTGGCCCTGAGAGGAGGCGCAGAGGGGCTTGGGGCTCAAGCAGCCAGCTCCCC 1827
Qy |||||
Db 501 ArgValThrAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThr 520
Qy |||||
Db 1828 AGGTGACAGAGCGGCTTGTACCCAGAGGAGCAGGAGCAGGAGCAGGAGCAGGAGCAGGAG 1887
Qy |||||
Db 521 GlyProGlnGluProThrSerAlaArgAla 530
Qy |||||

Db 1888 GGGCCTCAAGAGCCACCTCCGCCAGGGCC 1917

RESULT 7

ID ACC46130 standard; cDNA; 6577 BP.

XX ACC46130;

AC ACC46130;

DT 02-JUN-2003 (first entry)

XX Human dithp extracellular signalling protein-encoding cDNA.

Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;
cancer; cell proliferative disorder; autoimmune disorder;
inflammatory disorder; infection; hormonal disorder; metabolic disorder;
neurological disorder; gastrointestinal disorder; transport disorder;
connective tissue disorder; drug screening; proteome analysis;
gene therapy; antisense therapy; genotyping; transgenic animal; knock in;
disease model; toxicological testing; transcript imaging;
extracellular signalling; gene; ss.

XX Homo sapiens.

XX WO200297031-A2.

PN 05-DEC-2002.

PD 27-MAR-2002; 2002WO-US010056.

XX 28-MAR-2001; 2001US-0279619P.

XX 29-MAR-2001; 2001US-0280067P.

XX 29-MAR-2001; 2001US-0280068P.

XX 16-MAY-2001; 2001US-0291280P.

XX 17-MAY-2001; 2001US-0291829P.

XX 17-MAY-2001; 2001US-0291849P.

XX 19-JUN-2001; 2001US-0299428P.

XX 20-JUN-2001; 2001US-0299762P.

XX 20-JUN-2001; 2001US-0300001P.

XX (INCY-) INCYTE GENOMICS INC.

XX Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;

XX Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Ameshey SR;

XX Daughtery SC, Dam TC, Liu TP, Nguyen DA, Kleefeld Y, Gershin EH;

XX Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;

XX Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;

XX WPI; 2003-129518/12.

XX P-PSDB; ABR41186.

Novel human diagnostic and therapeutic polypeptide useful for identifying
test compound which specifically binds to a polypeptide encoded by human
diagnostic and therapeutic polynucleotide, and to induce antibodies.

Claim 2; SEQ ID NO 51; 591pp; English.

The invention relates to novel human diagnostic and therapeutic
polynucleotides designated dithp (ACC46080-ACC46749) and to their encoded
proteins (DITHP; ABR41136-ABR4182). The invention also relates to
polynucleotide sequences at least 90% identical to the dithp cDNA
sequences of the invention; recombinant vectors, host cells and
transgenic organisms comprising a dithp nucleic acid sequence; the
recombinant production of DITHP proteins; antibodies specific for DITHP
proteins; microarrays comprising dithp nucleic acid sequences; methods of
detecting dithp nucleotide and protein sequences; methods of screening
for compounds which specifically bind a DITHP protein; and methods of
assessing the toxicity of test compounds using a dithp hybridisation
probe. Dithp nucleic acid sequences and DITHP proteins may be used in the
diagnosis of a wide variety of conditions including cancer and other cell
proliferative disorders; autoimmune or inflammatory disorders; bacterial,
viral, fungal or parasitic infections; hormonal disorders; metabolic
disorders; neurological disorders; gastrointestinal disorders; transport
disorders; and connective tissue disorders. They may also be used to

CC screen for modulators of protein activity or gene expression. DITHP
CC proteins can additionally be used in analysis of the proteome of a tissue
CC or cell type and to induce antibodies. The dithp nucleic acids are
CC additionally useful in somatic or germline gene therapy of the disorders
CC mentioned above, as a source of antisense sequences, as a source of
CC probes and primers, in genotyping and identification of individuals, in
CC the generation of transgenic animal models of human disease or knock in
CC humanised animals, in toxicological testing, and in transcript imaging.
CC The present sequence represents a dithp cDNA encoding a DITHP protein
CC which has extracellular signalling activity. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 6577 BP; 1496 A; 1701 C; 1836 G; 1540 T; 0 U; 4 Other;

Alignment Scores: Length: 6577
Pred. No.: 7,29e-157 Matches: 529
Score: 2693.00 Conservatve: 0
Percent Similarity: 99.81% Mismatches: 1
Best Local Similarity: 99.81% Indels: 0
Query Match: 99.67% Gaps: 0
DB: 7

US-10-023-523-8 (1-530) x ACC46130 (1-6577)

QY	1	LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln	20
DB	328	AAAGAGAGCCAGGACAAACCGAAGAGGACCCGAGGAGCCAGGAGCGCCAGCCAG	387
QY	21	AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu	40
DB	388	CGCGCTCTGCTAGTAGAGCAGAGAGTCCCGGAGGAGCCAGGAGTCTCTGAGAGCGGAG	447
QY	41	GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer	60
DB	448	GGGGCTCAAGCCAGAGAGCGGTCTAGTCTGGGGCCCTCTGCTGATCTCTCTGAGGAGCTGAGC	507
QY	61	ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyProGly	80
DB	508	CGCAACTGGAAGACATCTAGACACATCTGTGTGACATTAACCGAGGGGGCCCGCGC	567
QY	81	GluAspGlyAlaGlnGlyGluProAlaGluProGluAlaGlnGlySerArgThrTyr	100
DB	568	GAGGATGGGGCACAGGTGAGCGGCTGAACCCGAGATGACAGAGAGTCCCGGACCTAT	627
QY	101	ValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGlyGluProSer	120
DB	628	GTGCAAGGAATGGGAGCGCTGAACCACTCCAGTCTCAATGGAGAGAGAACCTCTC	687
QY	121	LysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHis	140
DB	688	AAAGGGGATCCAAACACAGAGAGATCCGCGCAGAGTCAAGGTCGAGAGACCGAGCCAT	747
QY	141	ArgArgProGlnGluLysLysAlaLysGlyLysGlyGluIleThrLeuLeuMet	160
DB	748	CGAAGGGCCACAGAGAGAAAGCAAGGGTTTGGGAGGAGATCAGCTTGTCTGATG	807
QY	161	GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysTyr	180
DB	808	CAGACATTGATATCTCTGAGTACCCAGAGAGAGTGGCTGCTCTGTGCAAGAGATAT	867
QY	181	AlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLys	200
DB	868	GCTGAACCTGCTGGAGGAGCAGCGAATTCACAGAGGAGATGAAGCTCTTACAGAAAG	927
QY	201	GlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeu	220
DB	928	CAGAGCCAGCTGGTCCAAAGAGAGGACCACTTGGCGGTGAGCACACAGAGCGGCTCTG	987
QY	221	AlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLys	240
DB	988	GCCCGCAGCAAGCTTGAGAGCTATGCCGTGAGCTGACGCGGACCAACCGCTCCCTCAAG	1047

QY	241	GluGluGlyValGlnArgAlaArgGluGluGluGluLysArgLysGluValThrSerHis	260
DB	1048	GAGAGAGTGTGACGGGGCCCGGAGGAGGAGAGCGGAGGAGTGTGACCTCGCAC	1107
QY	261	PheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSer	280
DB	1108	TTCCAGGTGACACTGATGACATTCAGCTGAGATGAGACAGACCAATGAGCCAACTCC	1167
QY	281	LysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTyr	300
DB	1168	AAGTGTGCGCAAGAGAACTGAGAGCTGGCTGAGAGGCTCAGAGGCTGATGAGCAGTAT	1227
QY	301	GluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeu	320
DB	1228	GAGTGTGCGGAGGAGCATATCGCAAAAGTCTTCAACACAGAGACCTTACACAGCAGCTG	1287
QY	321	ValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGln	340
DB	1288	GTGGATGCCAAGCTCCAGCAGGCGGAGGATGCTAAAGAGGAGGAGAGCGGCCAG	1347
QY	341	ArgGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMet	360
DB	1348	CGGAGAGAGGATTTCTCTGAGAGGAGGAGTCCAGAGGAGTGTGTGAGCTGATG	1407
QY	361	LysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGlu	380
DB	1408	AAGCAGCAAGAGAGCCACCTGAGCAACAGCTTCCCTTATACACAGAGAGATTTGAGGAG	1467
QY	381	PheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGlu	400
DB	1468	TTCCAGAACACACTTTTCCAAAGCAGCAGGATATTCACACATTCAGAGCAGGATGAA	1527
QY	401	LysMetThrLysLysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArgTyr	420
DB	1528	AAGATGACTAAGAGATCAAGAGCTGAGAGAAAGAAACCACTATGTCCTCCCGTGG	1587
QY	421	GluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGlu	440
DB	1588	GAGAGCAGCAACAGGCGCTCTGAGATGCTGAGGAGAGAAACAGCTCCGGGATAAAGAA	1647
QY	441	LeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThr	460
DB	1648	CTGAGAGCGCTCGAGTAAATCAACGGCTGAGAGAGCTGTCCTCCGGCAGTCAAGACA	1707
QY	461	GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeu	480
DB	1708	GAGCGCAATGACCTGAAACAGAGGAGTACAGACCTGAGTGTGTGGCCAGGGCTCCCTC	1767
QY	481	ThrAspSerGlyProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerPro	500
DB	1768	ACTGACGTGGCCCTGAGAGGAGGAGCCAGGGGCTTGGGCTCAAGCAGCCAGCTCCCCC	1827
QY	501	ArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThr	520
DB	1828	AGCGTCAACAGAGCGCTTGTCTACCCAGAGCAGCCAGCAGAGCATCAGGCCAGACT	1887
QY	521	GlyProGlnGluProThrSerAlaArgAla	530
DB	1888	GGGCTTCAAGAGCCCACTCCCGCAGGCGC	1917

RESULT 8

AAV32836
ID AAV32836 standard; cDNA; 4722 BP.
XX
XX AAV32836;
AC
DT 09-NOV-1998 (first entry)
XX
DE Rabbit low density lipoprotein binding protein LBP-3 cDNA.
XX
KW Low density lipoprotein binding protein; LDL binding protein 2; LBP-2;
KW receptor; rabbit; atherosclerosis; diagnosis; therapy; vaccine; ss.


```

Db 1366 TGGGAGCAGCAGCAAGGCGCTTGATGGCTGAGGAGAAACACTCCGGGACAAA 1425
QY 440 GluLeuGluGlyLeuGlnValTylleGlnArgLeuGluLysLeuCysArgAlaLeuGln 459
Db 1426 GAGCTGAGAGGCGCTCAGGTGAAATCCAGCGCTGGAGAGCTGTGCCGGGCACTGCGAG 1485
QY 460 ThrGluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlnGlySer 479
Db 1486 ACAGAGCGCAATGACCTGAACAGAGGGTGCAGGACCTGAGTCCGGTGGCCAGGCGCCC 1545
QY 480 LeuThrAspSerGlyProGluArgArgPro----- 489
Db 1546 GTCTCCGAGCAGCGGTCTCTGAGCGGAGCGCCAGCGCCGCCACCACTCTCAAGAGAGGGGT 1605
QY 490 ---GluGlyProGlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyr 508
Db 1606 GTCGAGGCGCCCGGCGCTCAAGTACCCTCACTCCAAAGGGCCACAGAGCGTTCCTGCTGC 1665
QY 509 ProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAla 528
Db 1666 GCAGGTGCAACCCAGCAGCAGGATCATCAGGCGCCCGCCAGGAGCCCACTGCC 1725
QY 529 ArgAla 530
Db 1726 ACTGCC 1731

RESULT 9
AAH26491
ID AAH26491 standard; cDNA; 4722 BP.
XX
AC AAH26491;
XX
DT 12-NOV-2001 (first entry)
XX
DE Rabbit low density lipoprotein binding protein 3 (LBP-3) cDNA.
XX
KW Low density lipoprotein binding protein 3; LBP-3; LDL; rabbit;
KW atherosclerosis; antiarteriosclerotic; gene therapy; diagnosis; vaccine;
XX ss.
XX
OS Oryctolagus cuniculus.
XX
FH Key
FT Location/Qualifiers
FT 25..1734
FT CDS /*tag= a
XX
PN WO200164874-A2.
XX
PD 07-SEP-2001.
XX
PF 28-FEB-2001; 2001WO-US006356.
XX
PR 02-MAR-2000; 2000US-00517849.
XX
PR 14-JUL-2000; 2000US-00616289.
XX
PA (BOST-) BOSTON HEART FOUND INC.
XX
PI Lees AM, Lees RS, Law SW, Arjona AA;
XX
PI WPI: 2001-565505/63.
XX
PI P-PSDB; AAB82801.
XX
PT New isolated low density lipoprotein binding polypeptide for treating,
PT diagnosing and/or identifying therapeutic agents for atherosclerosis.
XX
PS Claim 5; Fig 14; 143pp; English.
XX
CC The present sequence is that of cDNA encoding novel rabbit low density
CC lipoprotein binding protein 3 (LBP-3, see AAB82801). The cDNA was
CC isolated following screening of a rabbit cDNA library for clones encoding
CC LBPs that bound to both native low density lipoprotein (LDL) and methyl
CC LDL. The invention provides claimed polynucleotides encoding novel

```

CC polypeptides which are capable of binding to native and methylated LDL,
 CC the isolated polypeptides, termed LBPs, and biologically active fragments
 CC and analogues of them, as well as expression vectors, cells and methods
 CC of producing the LBPs. Also claimed are methods for determining if an
 CC animal is at risk for atherosclerosis, methods for evaluating a cell
 CC for use in treating atherosclerosis, and methods for treating a cell
 CC having an abnormality in structure or metabolism of LBP. Pharmaceutical
 CC compositions comprising an LBP polypeptide or nucleic acid, and vaccine
 CC compositions, are also claimed

SQ Sequence 4722 BP; 997 A; 1332 C; 1430 G; 963 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5,75e-143 Length: 4722
 Score: 2465.50 Matches: 492
 Percent Similarity: 92.99% Conservative: 12
 Best Local Similarity: 90.77% Mismatches: 25
 Query Match: 91.25% Indels: 13
 DB: 5 Gaps: 3

US-10-023-523-8 (1-530) x AAH26491 (1-4722)

```

QY 1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyValaGlnGluArgProSerGln 20
Db 109 AAAGCAGCCCGGAGCAGCGCGAGGAGCGGAGCGGCGCGCGCGCGCGCGCG 168
QY 21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40
Db 169 CCGGCGCCCGCGGAGAAAGCGGAAGT---GCCAGCAGCAGCGCTCCCGGAGCGGAG 225
QY 41 GluValaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 60
Db 226 GGGGCTCAGCCAAACTGCTCAGCTGGGGCGCTCTGTGATGCTCTAGAGAGCTGAGC 285
QY 61 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGly 80
Db 286 CGCCAGTTGGAAGACATCTCTAGTACTGTGTGCACACACACACAGCGGCGCGCGGT 345
QY 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
Db 346 GAGGATGGGTCCAGGGTGAGCCCTGAACCTGAAGATGCAGAGAGTCTCGCGCTAT 405
QY 101 ValAlaArgAsnGlyGluProGluPro---ThrProValValTyrGlyGlyLysGluPro 119
Db 406 GTGCGAAGGAATGGGAGCGGAGCGGCGGAGCGGCGGAGTCAATGCGCAGAGGAGACC 465
QY 120 SerLysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAsp 139
Db 466 TCCAGCGCAGAGCGCGGCGCAGGAAGATCCGAGCGGATCGAGTCCGAGAGCCGAGAC 525
QY 140 HisArgArgProGlnGluLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeu 159
Db 526 CACCGAGGCCACAGAAAGAGAGGCGGAGGCTCTGGAAAGAGAGATCAGCTGCTG 585
QY 160 MetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLys 179
Db 586 ATGCAGACACTGAACACGCTGAGCACCCCGAGAGGAGAGCTGCGCGCTCTGTGAGAGAG 645
QY 180 TyrAlaGluLeuLeuGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLys 199
Db 646 TATCGGMACTGCTCGAGGAGGACCGGAATCTGCAGAGCAGATGAGTGTGTCAGAGAG 705
QY 200 LysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaVal 219
Db 706 AAGCAGAGCAGCTGCTGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 765
QY 220 LeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeu 239
Db 766 CTGGCGCGCAGCAGAGCTCGAGAGCGCTGTGCCGGGAGCTGCAGCGGACACACCGCTGCTC 825
QY 240 LysGluGluGlyValGlnArgAlaArgGluGluGluGluLysArgLysGluValThrSer 259
Db 826 AAGGAAGAAGGTGTGCAGCGAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 885

```

QY 260 HisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsn 279
 Db 886 CACTTCAGATGACGCTCAACGACATTCAGCTCAGATGAGCAGCACACAGCGCAAC 945
 QY 280 SerLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLeuLeuGluGln 299
 Db 946 TCCAGCTGCGCCAGGAGAACATGGAGCTGCGCGAGCGGCTCAAGAGCTGATTGAGCAG 1005
 QY 300 TyrGluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGln 319
 Db 1006 TACGAGCTGCGAGAGAGACATCTCAAAAGTCTTCAAAACACAGAGATCTGAGCAGCAG 1065
 QY 320 LeuValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluLysHis 339
 Db 1066 CTGCTGACCGCCAGCTTCCAGCAGGCGCCAGAGATCTGAGGAGGAGCAGGAGCGGCAC 1125
 QY 340 GlnArgGluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeu 359
 Db 1126 CAGCGGAGAGAGACTTCTCTCTGAAGGAGCGCGTGGAGTCCACAGGATGTGCGAGCTG 1185
 QY 360 MetLysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGlu 379
 Db 1186 ATGAAGCAACAGGAGAGCCACCTGAAGCAGCAGCTTGCCTATACACAGAGAGTTTGAG 1245
 QY 380 GluPheGlnAsnThrLeuSerLysSerSerGluValPheThrPheLysGlnGluMet 399
 Db 1246 GAGTTCAGAACACTCTTTCCTCAAAAGCAGCGAGGTGTTCAACACATTCAAACAGGAATG 1305
 QY 400 GluLysMetThrLysLysLysLysLysLeuGluLysGluThrThrMetTyrArgSerArg 419
 Db 1306 GAAAGATGACAAAGAGATCAAGAGCTGGAGAAAGACACCATGTACCGTTCCCGG 1365
 QY 420 TrpGluSerSerAsnLysAlaLeuLeuGluMetAlaGluLysThrValArgAspLys 439
 Db 1366 TGGAGAGCAGCAACAGGCGCTGTTGAGATGGCTGAGAGAAACACTCGCGCAGCAA 1425
 QY 440 GluLeuGluLysLeuGlnValLysLysLysLysLeuGluLysLeuCysArgAlaLeuGln 459
 Db 1426 GAGCTGAAGCGCTGAGGTGAAGAAATCCAGCGGCTGGAGAGCTGTGCGGSCACTGGAG 1485
 QY 460 ThrGluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySer 479
 Db 1486 ACAGAGCGCAATGACCTGACACAGAGGCTGAGGACCTGAGTGGCGGTGGCCAGGCCCC 1545
 QY 480 LeuThrAspSerGlyProGluArgArgPro----- 489
 Db 1546 GTCTCCGACAGCGGCTCTGAGCGGAGCGCCAGAGCGCCGCCACCTCCCAAGGAGCAGGGT 1605
 QY 490 ---GluGlyProGlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyr 508
 Db 1606 GTCGAGGCGCCCGGGCTCAAGTACCACTCTCCAGGCGCCACAGACGCTCTCTGCTGC 1665
 QY 509 ProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAla 528
 Db 1666 GCAGGTGACCCAGCAGAGGATCAGGCGCAGAGCGCCAGGCGCCCGCCAGGCGCCACTGCC 1725
 QY 529 ArgAla 530
 Db 1726 ACTGCC 1731

RESULT 10

ID AAS65026 standard; cDNA; 5284 BP.
 AC AAS65026;
 XX
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #830.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.
 OS
 XX WO200175067-A2.
 PN
 XX 11-OCT-2001.
 PD
 XX 30-MAR-2001; 2001WO-US008631.
 PF
 XX 31-MAR-2000; 2000US-00540217.
 PR
 XX 23-AUG-2000; 2000US-00649167.
 PR
 XX (HYSE-) HYSEQ INC.
 PA
 XX Drmanac RT, Liu C, Tang YT;
 PI WPI; 2001-639362/73.
 PT P-PSDB; ABG00839.
 DR
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 1; SEQ ID NO 830; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 5284 BP; 1370 A; 1419 C; 1627 G; 868 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,24e-135 Length: 5284
 Score: 2343.50 Matches: 483
 Percent Similarity: 85.66% Conservative: 7
 Best Local Similarity: 84.44% Mismatches: 22
 Query Match: 86.73% Indels: 61
 DB: 5 Gaps: 5

US-10-023-523-8 (1-530) x AAS65026 (1-5284)

QY 1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 20
 Db 502 AAAAGCAGCCAGGACCAACCGGAAGCAGGACCCGAGGAGGCCAGGAGCGCCAGCAG 561
 QY 21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40
 Db 562 GCGGCTCTCGAGTAGAAGCAGAGGTCCCGGACACGACGCGCTCTCGAGACCGGAG 621
 QY 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 60
 Db 622 GCGGCTCAAGCCAGACACGCTCAGTCTGGGGCGCTCTCGTATGTCTCTGAGGAGCTGAGC 681
 QY 61 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGly 80

682	Db	CGCCAACTGGAAGACATATCTGAGCACATATCTGTGTGGCAATAAACGAGGGGGGGCCCCGGC	741
81	Qy	GluAspGlyValaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr	100
742	Db	GAGGATGGGGCACAGGCTGAGCGGCTGAAACCCGAAGATGCAGAGAAGTCCCGGACCTAT	801
101	Qy	ValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGluLysGluProSer	120
802	Db	GTGGCAAGAAATGGGGAGCCTGAACCAATCCAGTAGTCAATGGAGAGAAGAAACCCCTCC	861
121	Qy	LysGlyAspProAsnThrGluLysLysAlaLysGlyLeu-----	140
862	Db	AAGGGGGATCCAAACACAGAGAGATCCGGCAGAGTGCAGAGGTCCGAGACCCGAGACCAT	921
141	Qy	ArgArgProGlnGluLysLysAlaLysGlyLeu-----	152
922	Db	CGAAGGCCACAGGAGAAAGAAAAGCCAAAGGTTTGGGTGAGCAGAGGGCGGCTTTTGT	981
153	Qy	-----GlyLysGluLysLeuMetGlnThrLeuAsnThrLeuSerThrProGlu	170
982	Db	GAAGCTGGGAAGGAGATCACCTTCTGTATGACAGACATTGGAATATCTCTGAGTACCCACAG	1041
171	Qy	GluLysLeuAlaLeuLysLysLysTyrAlaGluLeuLeuGluGluHisArgAsnSer	190
1042	Db	GAGAAGCTGGCTCTCTGTGCAAGAGTATGCTGAACTGCTGGAGAGGACCCGGGAATCA	1101
191	Qy	GlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHis	210
1102	Db	CAGAAGCAGATGAAGCTCTACAGAAAAGCAGAGCCAGCTGGTGTCAAGAGAAGACAC	1161
211	Qy	LeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArg	230
1162	Db	CTGGCGGTGAGCACAGCAAGGCCGTCTGTGCCCGCAGCAAGCTTGAGAGCCTATGCCGT	1221
231	Qy	GluLeuGlnArgHisAsnArgSerLeuLys-----	240
1222	Db	GAGCTGAGCGGCACAAACCGCTCCCTCAAGCTGGGGGACCCCTTGTGAGGTAGCAGTGGAT	1281
241	Qy	-----GluGluGlyValGlnArgAlaArgGluGlu	250
1282	Db	GATGGCTCTCGGGCTGACTCTTTTGGCCAGAGAAGGTGTGACAGGGGGCCCCGGGAGGAG	1341
251	Qy	GluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeu	270
1342	Db	GAGGAGAAGCGCAAGGAGGTGACCTCGCACTCCAGGTGACACTGAAATGACATTCAGCTG	1401
271	Qy	GlnMetGluGlnHisAsnGluLysArgSerLysLeuArgGlnGluAsnMetGluLeuAla	290
1402	Db	CAGATGAAACAGACACAATGACGCGCACTCAAGCTGCGCAAGAGAACAATGGAGCTGGCT	1461
291	Qy	GluArgLeuLysLysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAspLysVal	310
1462	Db	GAGAGGCTCAAGAAGCTGATTGAGCAGTATGAGCTGCGCGAGAGGCAATTCGACAAGTTC	1521
311	Qy	PheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGlu	330
1522	Db	TTCAACACAAAGGACCTTACAAACAGCAGCTGGTGGATGCCAAGCTCCAGCAGGCCCCAGGAG	1581
331	Qy	MetLeuLysGluAlaGluGluArgHisGlnArgGluLysAspPheLeuLysGluAla	350
1582	Db	ATGCTAAGAGGACGAGAGAGCGGACCGAGCGGAGAGAGATTTCTCTGAAAGAGGCA	1641
351	Qy	ValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlnThrHisLeuLysGlnGln	370
1642	Db	GTAGAGTCCAGAGGATGTGTGAGCTGATCAAGCAGCAAGAGAGACCCACCTGAAGCAACAG	1701
371	Qy	LeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrIeuSerLysSerGlu	390
1702	Db	CTTGCCCTATACACAGAGAATTTGAGGAGTTCAGAACACACTTTTCCAAAAGACGCGAG	1761
391	Qy	ValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysIleLysLysLeuGlu	410

Db	1762	GTATTCCACCATTCACACGAGAGATGAAAGATGACTAAGAAGATCAAGAAGCTGGAG	1821
QY	411	LysGluThrThrMetTyrArgSerArgTrpGluSerAsnLysAlaLeuLeuGluMet	430
Db	1822	AAAGAAACCATGATCCGGTCCGGTGGGAGAGCAGCAACAAGGCCCTGCTTGAGATG	1881
QY	431	AlaGluGluLysThrValArgAspLysGluLeuGluGlyLeuGlnValLysIleGlnArg	450
Db	1892	GCTGAGGAGAAACAGTCCGGGATAAGACTGGAGGCGCTGCAGGTAAATAATCCAAACGG	1941
QY	451	LeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgValGln	470
Db	1942	CTGAGAGAAGCTGTGCCGGCACTGCAGACAGAGCGCAATGACCTGTAACAAGAGTCTGGG	2001
QY	471	AspLeu-----SerAlaGlyGlyGln	477
Db	2002	CCCTGGTTATGAAAAGTTCGGAATCTCTTTACCTTGTGGAGTGTCTGAGTGTT	2061
QY	478	GlySerLeuThrAspSerGlyProGluArgArgProGluGlyProGlyAlaGlnAlaPro	497
Db	2062	GGAAAGTACCCAGGAAGA-AGCCCTGAGCAGGTGCCCTCAGGA-----GCAGTGCCCC	2111
QY	498	SerSerProArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAlaSer	517
Db	2112	ATGCTCCCMCATCAGCCAA-----	2132
QY	518	GlyGlnThrGlyProGlnGluProThrSerAlaArg	529
Db	2133	---GAGGCCCAACCCAGGAAGCCACTCCTCGCCCGG	2165
RESULT 11			
AAKS1890			
ID	AAKS1890 standard; cDNA; 2188 BP.		
XX	AAKS1890;		
AC	06-NOV-2001 (first entry)		
DT	Human polynucleotide SEQ ID NO 435.		
DE	Human; cytokine; cell proliferation; cell differentiation; gene therapy;		
KW	vaccine; peptide therapy; stem cell growth factor; haematopoiesis;		
KW	tissue growth factor; immunomodulatory; cancer; leukaemia;		
KW	nervous system disorder; arthritis; inflammation; ss.		
OS	Homo sapiens.		

RESULT 11

AAK51890

ID AAK51890 standard; cDNA; 2188 BP.

AC AAK51890:

06-NOV-2001 (first entry)

Human polynucleotide SEQ ID NO 435

Human: cytokine, cell, evolution,

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation; ss.
 KW

Homo sapiens.

WO200157190-A2

09-AUG-2001

XX
PF
05-FEB-2001: 2001WO-IIS004098XX
PP 03-FEB-2000. 2000US-00486014

PR 27-APR-2000; 2000US-00560875.

PR 19-JUL-2000; 2000US-00620325.

15-SEP-2000; 2000US-00663561.

PR 30-NOV-2000; 2000US-00728422.

PA (HYSE-) HYSEQ INC.

PI. Tang YT, Liu C, Drmanac RT,

PI Xue AJ, Yang Y, Wejhrman T,

DR WPI; 2001-476283/51.

XX

in diagnosis and gene therapy.

XX Claim 1; Page 1605-1607; 6221pp; English.
 PS The invention relates to polynucleotides (AAK51456-AAK53435) and the
 XX encoded polypeptides (AAM78323-AAM80302) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
 CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
 CC sequence listing were missing at the time of publication
 XX
 SQ Sequence 2188 BP; 541 A; 559 C; 669 G; 418 T; 0 U; 1 Other;

Alignment Scores:
 Pred. No.: 2,08e-99 Length: 2188
 Score: 1754.50 Matches: 349
 Percent Similarity: 98.31% Conservatives: 0
 Best Local Similarity: 98.31% Mismatches: 5
 Query Match: 64.93% Indels: 1
 DB: 4 Gaps: 1

US-10-023-523-8 (1-530) x AAK51890 (1-2188)

QY 177 CysLysLysTyr---AlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLys 195
 DB 358 TGCCCTCCCTATCCCTGCTCCCTGAGGAGCAGCCGAAATTCACAGAGCAGATGAAG 417
 QY 196 LeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHis 215
 DB 418 CTCTACAGAAAGAGAGCCAGCTGCTGACAGAGAGGACCCACTGCTGCGGTGAGCAC 477
 QY 216 SerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHis 235
 DB 478 AGCAAGCCGCTCTGCGCCGAGCAAGCTTGAGAGCCATATGCTGAGCTGAGCGGCAC 537
 QY 236 AsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGluGluGluLysArgLys 255
 DB 538 AACCGCTCCCTCAGGAGAGAGAGGTGTGACGCGGCGCCGAGGAGGAGAGAGCCAG 597
 QY 256 GluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHis 275
 DB 598 GAGGTGACCTCGCACTTCCAGGTGACACTGAATGACATTCAGCTGCAGATGGAACAGCAC 657
 QY 276 AsnGluArgAsnSerLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLys 295
 DB 658 AATGAGCGCAACTCCCAAGCTGCGCCCAAGAGAACATGAGCTGCTGAGAGCTCAAGAAG 717
 QY 296 LeuIleGluGlnTyrGluLeuArgGluGluHisIleAspLysValPheLysHisLysAsp 315
 DB 718 CTGATTGACGAGTATGAGCTGCGGAGGAGCATATCGACAAAGTCTTCAACACAGAGGAC 777
 QY 316 LeuGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAla 335
 DB 778 CTACAACAGAGCTGTGTGGATGCCAAGCTCCAGAGCCCGAGAGATGCTTAAGAGGAGCA 837
 QY 336 GluGluArgHisGlnArgGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArg 355
 DB 838 GAAGAGCGGACCCAGCGGAGAGAGGATTTCTCTGAAAGAGCGAGTAGTGCTCCAGAGG 897
 QY 356 MetCysGluLeuMetLysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuThr 375
 DB 898 ATGTGTGAGCTGTATGAAGCAGCAGAGACCCACTGAAGCAACAGCTTGCCCTATACACA 957
 QY 376 GluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluValPheThrPhe 395
 DB 958 GAGAAGTTTGAGGAGTTTCCAGAACACACTTTCCAAAGAGCGAGGATTTCCACCATTC 1017

QY 396 LysGlnGluMetGluLysMetThrLysLysLysLysLysLeuGluLysGluThrThrMet 415
 DB 1018 AAGCAGGAGATGAAAGATGACTAAGAGATCAAGAGCTGAGAGAAAGAAACCCACCATG 1077
 QY 416 TyrArgSerArgTrpGluSerSerAsnLysAlaLeuLeuGluMetAlaGluLysThr 435
 DB 1078 TACCGTCCCGTGGGAGAGCAGCAACAGGCCCTGCTTGAGATGGCTGAGGAGAAACA 1137
 QY 436 ValArgAspLysGluLeuGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 455
 DB 1138 GTCCCGGATAAAGAACTGGAGGCGCTGCAGGTAAATAATCCACGCTGAGAGAGTGTGC 1197
 QY 456 ArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGly 475
 DB 1198 CGGGCACTGCACAGAGCGCAATGACCTGAACACAGAGGGTACAGAGCTGAGTCTGGT 1257
 QY 476 GlyGlnGlySerLeuThrAspSerGlyProGluArgArgProGluGlyProGlyAlaGln 495
 DB 1258 GGCCAGGGCTCCCTCACTGACAGTGGCCCTGAGAGAGGCCACAGAGGGGCTGGGCTCAA 1317
 QY 496 AlaProSerSerProArgValThrGluAlaProCysTyrProGlyAlaProSerThrGlu 515
 DB 1318 GCACCCAGCTCCCGAGGTACAGAGCGCTTGTACCCAGGAGCACCAGGACACAGAA 1377
 QY 516 AlaSerGlyGlnThrGlyProGlnGluProThrSerAlaAlaArgAla 530
 DB 1378 GCATCAGGCGCAGACTGGGCTCAAGAGCCCACTCCGCCAGGGCC 1422

RESULT 12

ABZ34807/C
 ID ABZ34807 standard; cDNA; 1793 BP.
 XX
 AC ABZ34807;
 DT 04-FEB-2003 (first entry)
 XX
 DE Coding sequence SEQ ID 165, upregulated in osteogenesis.
 XX
 KW Osteopathic; osteogenesis modulator; gene therapy; osteogenesis;
 KW osteoporosis; bone disease; downregulator; human; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200281745-A2.
 XX
 PD 17-OCT-2002.
 XX
 PF 05-APR-2002; 2002WO-IB002211.
 XX
 PR 05-APR-2001; 2001US-0281400P.
 XX
 PA (AVET) AVENTIS PHARMA SA.
 XX
 PI Garcia T, Roman Roman S, Barton R, Call K, Theilhaber J;
 PI Connolly T, Jackson A, Bushnell SE, Rawadi G;
 XX
 DR WPI; 2003-058567/05.

XX Novel isolated nucleic acid upregulated/downregulated in osteogenesis,
 PT useful for bone disease therapy in subject.
 XX
 PS Claim 26; Page 165-166; 237pp; English.
 CC The present invention relates to novel nucleotide sequences, which are
 CC differentially expressed in models of osteogenesis upon being put in
 CC contact with a stimulator of osteogenesis. The present sequence is one
 CC such sequence. This sequence can be used for diagnosing osteoporosis/bone
 CC disease in a patient, promoting osteogenesis and/or preventing
 CC osteoporosis/bone disease. The present sequence encodes a secreted
 CC protein
 XX
 SQ Sequence 1793 BP; 339 A; 546 C; 444 G; 464 T; 0 U; 0 Other;

[illegible]

CC Sequences given in the disclosure of the present invention do not match
 CC up with their corresponding SEQ ID NO: sequences given in the sequence
 CC listing

XX Sequence 1793 BP; 339 A; 546 C; 444 G; 464 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,15e-95 Length: 1793
 Score: 1688.00 Matches: 339
 Percent Similarity: 99.12% Conservativity: 0
 Best Local Similarity: 99.12% Mismatches: 3
 Query Match: 62.47% Indels: 1
 DB: 3 Gaps: 0

US-10-023-523-8 (1-530) x AAA35218 (1-1793)

Qy 189 AsnSerGlnLysGlnMetLysLeuGlnLysLysGlnSerGlnLeuValGlnGluLys 208
 Db 1791 AATTACAGAGCAGATGAAGCTCTTACAGAAAAAGCAGAGCAGCTGTGTGCAAGAGAAG 1732
 Qy 209 AspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeu 228
 Db 1731 GACCACTGCGGTGAGCAGCAGCAAGCC-GTCTGCGCCGCGCAGAGCTTGAGAGCCTA 1673
 Qy 229 CysArgLeuGlnArgHisAsnArgSerLysLeuGluGluValGlnArgAlaArg 248
 Db 1672 TGCGGTGAGTGCAGCGGCACAAACCGCTCCCTCAGGAAAGAGGTGTGCGAGCGGCCCGG 1613
 Qy 249 GluGluGluGluLysArgLysGluValThrSerHisPheGluValThrLeuAsnAspLe 268
 Db 1612 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1553
 Qy 269 GlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGlnAsnMetGlu 288
 Db 1552 CAGCTGCAGATGGAACAGCAGCAATGAGCGCACTCCAGCTGCGCCAGAGAAATGCGAG 1493
 Qy 289 LeuAlaGluArgLeuLysLysLeuLysGluGlnTyrGluLeuArgGluHisLeuAsp 308
 Db 1492 CTGGCTGAGAGGCTCAAGAACTGATGAGCTGATGAGCTGCGGAGGAGGAGGAGGAGGAG 1433
 Qy 309 LysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAla 328
 Db 1432 AAAGTCTTCAACACAGAGGACCTTACACAGCAGCTGCTGATGCCAAGCTCCAGCAGGCC 1373
 Qy 329 GlnGluMetLysGlnAlaGluArgHisGlnArgGluLysAspPheLeuLysLys 348
 Db 1372 CAGGAGTCTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1313
 Qy 349 GluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlnThrHisLeuLys 368
 Db 1312 GAGCAGTAGAGTCCAGAGGAGTGTGAGCTGATGAAGCAGCAGAGACCCACCTGAAG 1253
 Qy 369 GlnGluLeuAlaLeuTyrThrGluLysPheGluGluPheGluAsnThrLeuSerLysSer 388
 Db 1252 CAACAGCTTCCCTATACAGAGAGAGTTGAGGAGTTCCAGAACACATTTTCCAAAAGC 1193
 Qy 389 SerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLys 408
 Db 1192 AGCAGGATTTACCACTTCAAGCAGGAGAGTGAAGATGATGAAGCAGCAGAGATCAAGAAG 1133
 Qy 409 LeuGluLysGluThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeu 428
 Db 1132 CTGGAGAAAGAACCACTGATACCCGCTCCGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1073
 Qy 429 GluMetAlaGluLysThrValArgAspLysGluLeuGluGluLysValLysLysLys 448
 Db 1072 GAGATGGCTGAGGAGAGAAACAGTCCCGGATAAAGAACTGAGGAGGCTGCGAGTAAATATC 1013
 Qy 449 GlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuLysArg 468
 Db 1012 CAACGGCTGAGAGAGCTGTGCGCGGACCTGCGAGCAGAGCGCAATGACCTGAACAGAGG 953
 Qy 469 ValGlnAspLeuSerAlaGlyGlnGlySerLeuThrAspSerGlyProGluArgArg 488

Db 952 GTACAGAGCTGAGTCTGTGTGCGCCAGGCTCCCTCACTGACAGTGGCCCTGAGAGAGG 893
 Qy 489 ProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyr 508
 Db 892 CCAGAGGGGCTGGGCTCAAGCACCAGCTCCCGGCTCAGAGAGGCGCTTGCTAC 833
 Qy 509 ProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGlnProThrSerAla 528
 Db 832 CCAGGAGACCGAGCAGAGCATCAGGCCAGAGCTGGGCTCAAGAGCCCACTCCGCC 773
 Qy 529 ArgAla 530
 Db 772 AGGCGC 767

RESULT 14

AAA35217/c

ID AAA35217 standard; DNA; 1793 BP.

XX AAA35217;

AC AAA35217;

DT 28-JUL-2000 (first entry)

XX Human adenosine receptor related polynucleotide 2nd SEQ ID NO:91.

DE Human; adenosine receptor; low adenosine antisense oligonucleotide;
 KW phosphorothioate; impaired respiration; inflammation; allergy;
 KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
 KW antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;
 KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
 KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;
 KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
 KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.

XX Homo sapiens.

XX WO200009525-A2.

XX 24-FEB-2000.

XX 03-AUG-1999; 99WO-US017712.

XX 03-AUG-1998; 98US-0095212P.

XX (UYEC-) UNIV EAST CAROLINA.

XX Nyce JW;

XX WPI; 2000-205971/18.

XX New antisense oligonucleotides useful for treating e.g. pulmonary
 vasoconstriction, inflammation, allergies, asthma, hypertension,
 bronchitis, emphysema, respiratory distress syndrome, ischemia or
 cancers.

XX Disclosure; Page 1259; 1343pp; English.

XX The present invention describes a new composition comprising an antisense
 oligonucleotide (ON) with low adenosine (up to 15%), which targets
 nucleic acids involved in bronchoconstriction, allergies, and/or
 inflammation. The ON can have antiinflammatory, antiallergic,
 antiasthmatic, cytostatic and analgesic activities. The compositions are
 useful for the treatment of diseases associated with inflammation,
 impaired airways, including lung disease and diseases whose secondary
 effects afflict the lungs of a subject. They can be used for treating
 e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma,
 COPD, respiratory distress syndrome, chronic obstructive
 pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
 carcinomas, and cancers which may metastasize to the lungs, including
 breast and prostate cancer. The reduction of the adenosine content of the
 ONs reduces side effects. The A-containing ONs break down with the
 release of deoxyadenosine which activates adenosine receptors causing

189	QY	AsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLys	208
1791	Db	AATTACAGAGCAGATGAAGCTCTACAGAAAAGCAGAGCAGCTGGTCAAGAGAAG	1732
209	QY	AspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerIeu	228
1731	Db	GACCACCTTGCCTGGGTGAGCACAGCAAGCC-GTCTTGGCCCGCAGCAGCTTGAGAGCCTA	1673
229	QY	CysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArg	248
1672	Db	TGCCGTGAGCTGCACGGGCACACCGCTCCCTCAGGAAAGAGGTGTGCACGGGCGCCGG	1613
249	QY	GluGluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIle	268
1612	Db	GAGGAGGAGGAGAAGCCCAAGGAGGTGACCTCGCACTTCCAGGTGACACTGAATGACATT	1553
269	QY	GlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGlnGluAsnMetGlu	288
1552	Db	CAGCTGCAGATGGAACAGACAAATGAGCCCACTCCAAGCTGCGCCCAAGAGAACAATGGAG	1493
289	QY	LeuAlaGluArgLeuLysLysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAsp	308
1492	Db	CTGGCTCAGAGGGCTCAAGAAGCTGATTGAGCAGTATGAGCTGCCGCGAGGAGCATATGCAC	1433
309	QY	LysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAla	328
1432	Db	AAAGTCTTCAACACAAAGGACCTTCAACAGCAGCTGGTGGATGCCAGCTCCAGCAGGCG	1373
329	QY	GlnGluMetLeuLysGluAlaGluGluArgHisGlnArgGluLysAspPheLeuLeuLys	348
1372	Db	CAGGAGTGTCTAAGGAGGAGAAAGAGCGGACACAGCGGAGAGAGGATTTCCTCTGAAA	1313
349	QY	GluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLeuLys	368
1312	Db	GAGCAGTAGAGTCCACAGAGGATGTGTGAGCTGATGAAGCAGCAGAGACCCACCTGAG	1253
369	QY	GlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSer	388
1252	Db	CAACAGCTTGGCCCTATACACAGAGAAGTTTGAAGAGTTCCAGAACACACATTTCCAAAAGC	1193
389	QY	SerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysIleLysLys	408
1192	Db	AGCGAGGTATTCCACCATTCGAACGAGAGATGGAAGAAGATGACTAAGAGATCAAGAAG	1133
409	QY	LeuGluLysGluThrThrMetTyrArgSerArgTrpGluSerSerAsnLysAlaLeuLeu	428
1132	Db	CTGGAGAAAGAAACACCATGTACCGTCCCGTGGGAGAGCAGCAACAGGCCCTGCTT	1073
429	QY	GluMetAlaGluGluLysThrValArgAspLysGlnLeuGluGlyLeuGlnValLysIle	448
1072	Db	GAGATGGCTGAGGAGAAACAGTCCGGGATAAAGACTGGAGGGCTCGAGGTAAAAATC	1013

activating peptide factors and transmitters, transcription factors, immunoglobulins and antibodies, antibody receptors, cytokines and chemokines, endogenously produced specific and non-specific enzymes, binding proteins, adhesion molecules and their receptors, cytokine and chemokine receptors, adenosine receptors, bradykinin receptors, central nervous system (CNS) and peripheral nervous and non-nervous system receptors, CNS and peripheral nervous and non-nervous system peptide transmitters, defensins, growth factors, vasoactive peptides and receptors, binding proteins and malignancy associated proteins. The antisense oligonucleotides may be used in this way to treat disorders including respiratory obstruction (especially pulmonary obstruction and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or surfactant hypoproduction which are associated with a disease or condition selected from pulmonary vasoconstriction, inflammation, allergies, asthma, impeded respiration, respiratory distress syndrome (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), pulmonary transplantation rejection, pulmonary infections, bronchitis, and/or cancer. AAF18434 to AAF21543 represent human polynucleotide fragments and antisense oligonucleotides used in the exemplification of the present invention

Sequence 1793 BP; 339 A; 546 C; 444 G; 464 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,15e-95 Length: 1793
Score: 1688.00 Matches: 339
Percent Similarity: 99.12% Conservative: 0
Best Local Similarity: 99.12% Mismatches: 3
Query Match: 62.47% Indels: 1
DB: 3 Gaps: 0

US-10-023-523-8 (1-530) x AAF21340 (1-1793)

Qy 189 AsnSerGlnLysGlnMetLysLeuGlnLysGlnLysGlnLysGlnLysValGlnLys 208
Db 1791 AATTCACAGAGCAGATGAAGCTCTTACAGAAAAAGCAGAGCAGCTGTCACAGAGAAG 1732
Qy 209 AspHisLeuArgGlyCysHisSerLysAlaValLeuAlaArgSerLysLeuSerLeu 228
Db 1731 GACCACCTCGCGGTGAGCAGCAGAACCC-GTCTGCGCCGCGCAGCAGCTTGAGAGCTTA 1673
Qy 229 CysArgGluLeuGlnArgHisAsnArgSerLysLeuLysGluGlyValGlnArgAlaArg 248
Db 1672 TCCGTGAGCTCACCGGCACACCGCTCTCTCAAGGAAGAGGTGTGACGGGCCCGG 1613
Qy 249 GluGluGluGlyLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspLe 268
Db 1612 GAGGAGGAGGAGAGCGCAAGGAGGTGACCTCGCACTTCCAGGTGACACTGAATGACATT 1553
Qy 269 GlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGlu 288
Db 1552 CAGCTGCAGATGGAACAGCAGACATGAGCGCAACTCCAGCTCGCCCAAGAGACATGGAG 1493
Qy 289 LeuAlaGluArgLeuLysLysLeuLeuGluGlnTyrGluLeuArgGluGluHisLeasp 308
Db 1492 CTGGCTGAGAGGCTCAAGAAGCTGATTGAGCAGTATGAGCTGCGCAGGAGCATATCGAC 1433
Qy 309 LysValPheLysHisLysAspLeuGlnGlnLeuValAspAlaLysLeuGlnGlnAla 328
Db 1432 AAGTCTTCAACACAGAGACCTTACACAGCAGCTGTGTGATGCCAAGCTCCAGCAGGCC 1373
Qy 329 GlnGluMetLeuLysGluAlaGluGluArgHisGlnArgGluLysAspPheLeuLeuLys 348
Db 1372 CAGGAGATGCTAAAGGAGGAGCAGAGAGCGCACCGCAGCGGAGAGATTCTCTCTGAAA 1313
Qy 349 GluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnThrHisLeuLys 368
Db 1312 GAGGCACTGAGTCCAGAGGATGTGTGAGCTGATGAGCAGCAGCAGAGACCCACCTGAAG 1253
Qy 369 GlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSer 388
Db 1252 CAACAGCTTGCCTTATACACAGAGAGATTGAGAGTTCCAGACACACTTTCGAAAGC 1193

Qy 389 SerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLys 408
Db 1192 AGCGAGGTAATTCACCATTCAGAGCAGGAGATGAAAAGATGACTAAGAGATCAAGAAG 1133
Qy 409 LeuGluLysGluThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeu 428
Db 1132 CTGAGAAAGAAACACCACTGACCGGTCCCGTGGGAGAGCAGCAACAGGCGCTGCTT 1073
Qy 429 GluMetAlaGluGluLysThrValArgAspLysGluLeuGluGluLysLysLysLys 448
Db 1072 GAGATGCTCAGGAGAAAACAGTCCGGGATAAAGAACTCGAGGCGCTCCAGGTAAATAATC 1013
Qy 449 GlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArg 468
Db 1012 CAACGGCTGAGAGAGCTGTCCCGGGCACTGCACAGAGCGCAATGACCTGACACAGAGG 953
Qy 469 ValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSerGlyProGluArgArg 488
Db 952 GTACAGGACCTGAGTGTGCTGGCCAGGGCTCCCTCACTGACAGTGGCCCTGAGAGGAGG 893
Qy 489 ProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyr 508
Db 892 CCAGAGGGGCTCGGGCTCAAGCACCCAGCTCCCGCAGGTCCAGAGGCGCTTGTCTAC 833
Qy 509 ProGluAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAla 528
Db 832 CCAGAGCAGCAGCAGCAGAGAGATCAGGCCAGCTGAGGCGCTCAGAGCCACCTCCGCC 773
Qy 529 ArgAla 530
Db 772 AGGCGC 767

RESULT 16
AAF21339/c
ID AAF21339 standard; DNA; 1793 BP.
XX
AC AAF21339;
DT 14-MAR-2001 (first entry)
XX
DE Human low adenosine antisense oligonucleotide related sequence #2906.
KW Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
KW human; airway disorder; bronchoconstriction; lung inflammation;
KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;
KW immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
KW respiratory obstruction; pulmonary obstruction; impeded respiration;
KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;
KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
KW cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200062736-A2.
XX
PD 26-OCT-2000.
XX
PF 24-MAR-2000; 2000WO-US008020.
XX
PR 06-APR-1999; 99US-0127958P.
XX
PA (UYEC-) UNIV EAST CAROLINA.
XX
PI (NYCE/) NYCE J W.
XX
DR WPI; 2000-679539/66.
XX
PT Low adenosine (A) content antisense oligonucleotides which do not trigger
PT adenosine receptors during metabolism, useful e.g. for treating cancers

PT and respiratory obstructions.

XX Disclosure; Page 1343; 1592pp; English.

CC The present invention describes low adenosine (A) content antisense oligonucleotides and compositions (I) comprising them. In the antisense oligonucleotides the A is replaced by a 'Universal' or alternative base. CC (1) can have respiratory, bronchodilator, antiinflammatory, analgesic, CC immunosuppressive, antiasthmatic, hypotensive and cytostatic activities. CC The antisense oligonucleotides and (I) can be used to down-regulate the CC expression and/or activity of target polypeptides associated with the CC lung/respiratory disorders and malignancies, such as stimulating and CC activating peptide factors and transmitters, transcription factors, CC immunoglobulins and antibodies, antibody receptors, cytokines and CC chemokines, endogenously produced specific and non-specific enzymes, CC binding proteins, adhesion molecules and their receptors, cytokine and CC bradykinin receptors, adenosine receptors, bradykinin receptors, central CC nervous system (CNS) and peripheral nervous and non-nervous system CC receptors, CNS and peripheral nervous and non-nervous system peptide CC transmitters, defensins, growth factors, vasoactive peptides and CC receptors, binding proteins and malignancy associated proteins. The CC antisense oligonucleotides may be used in this way to treat disorders CC including respiratory obstruction (especially pulmonary obstruction CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or CC surfactant hypoproduction which are associated with a disease or CC condition selected from pulmonary vasoconstriction, inflammation, CC allergies, asthma, impeded respiration, respiratory distress syndrome CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD), CC pulmonary transplantation rejection, pulmonary infections, bronchitis, CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide CC fragments and antisense oligonucleotides used in the exemplification of CC the present invention

XX Sequence 1793 BP; 339 A; 546 C; 444 G; 464 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,15e-95 Length: 1793
Score: 1688.00 Matches: 339
Percent Similarity: 99.12% Conservative: 0
Best Local Similarity: 99.12% Mismatches: 3
Query Match: 62.47% Indels: 1
DB: 3 Gaps: 0

US-10-023-523-8 (1-530) x AAF21339 (1-1793)

QY 189 AsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLys 208
DB 1791 AATTCAGAGGAGGATGAGCTTCTACAGAAAAGCAGAGCGCGCTGGTGCAAGAGAG 1732
QY 209 AspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeu 228
DB 1731 GACCACCTGCGCGTGAGCACAGCAAGCC-GTCCTGCGCGCGAGCAAGCTTGAGAGCCTA 1673
QY 229 CysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGlyValGlnArgAlaArg 248
DB 1672 TGGCGGTGAGCTGACGGGCACAAACCCCTCCCTCAAGAGAAGGTGTGACGGCGCCGG 1613
QY 249 GluGluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIle 268
DB 1612 GAGGAGGAGGAGAGCGCAAGGAGGTGACCTCGACCTTCCAGGTGACACTGAATGACATT 1553
QY 269 GlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGlu 288
DB 1552 CAGCTGCATGATGAACAGACATATGAGCGCAACTCCAAAGCTGCGCAAGAGAACATGAG 1493
QY 289 LeuAlaGluArgLeuLysLysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAsp 308
DB 1492 CTGGCTGAGAGGCTCAAGAGCTGATTGACAGCTATGACCTGCGGAGGAGCATATCGAC 1433
QY 309 LysValPheLysHisLysAspLeuGlnGlnLeuValAspAlaLysLeuGlnGlnAla 328
DB 1432 AAAGTCTTCAACACAAAGGACCTTACAAACAGCAGCTGGTGGATGCCAAGCTCCAGCAGGCC 1373

QY 329 GlnGluMetLeuLysGluAlaGluGluArgHisGlnArgGluLysAspPheLeuLeuLys 348
DB 1372 CAGGAGATGCTAAAGAGGAGGAGAGCGGCACAGCGGAGAGGATTTTCTCTCTGAAA 1313
QY 349 GluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnThrHisLeuLys 368
DB 1312 CAGCAGTAGAGTCCAGAGAGTGTGTAGCTGATGAAGCAGCAGAGAGACCCACCTGAAG 1253
QY 369 GlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSer 388
DB 1252 CAACAGCTTGCCCTATACACAGAGAAGTTTCAGAGATTCAGAAACACACACTTCCAAAGC 1193
QY 389 SerGluValPheThrThrPheLysGlnGluMetCysLysMetThrLysLysIleLysLys 408
DB 1192 AGCGAGGTATTCCACCATTCAGCAGAGATGGAAGAAGATGACTAAGAAGATCAAGAAG 1133
QY 409 LeuGluLysGluThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeu 428
DB 1132 CTGGAGAAAGAACCCACCATGTACCGGTCCCGGTGGGAGAGCAGCAACAAGGCCCTGCTT 1073
QY 429 GluMetAlaGluGluLysThrValArgAspLysGluLeuGluGluLysValLysIle 448
DB 1072 GAGATGCTGAGGAGAGAAACAGTCCCGGATAAAGAACTGGAGGGCCTGCAGGTAAAAATC 1013
QY 449 GlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArg 468
DB 1012 CAACGGCTGGAGAGCTGTGCGGGCACTGCAGACAGAGCGCAATGACCTGAACAAGAGG 953
QY 469 ValGlnAspLeuSerAlaGlyGlnGlySerLeuThrAspSerGlyProGluArgArg 488
DB 952 GTACAGGACCTGAGTGTGTGGTGGCAGGGTCCCTCACTCACAGTGGCCCTTGAGAGAGG 893
QY 489 ProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyr 508
DB 892 CCAGAGGGCGCTGGGGCTCAAGCACCAGCTCCCCAGGGTCAAGAGGCCCTTGCTAC 833
QY 509 ProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAla 528
DB 832 CCAGGAGCAGCAGCAGCAGAGCATCAGGCGCAGACTGGGCTCAAGAGGCCACCTCCGCC 773
QY 529 ArgAla 530
DB 772 AGGGCC 767
RESULT 17
ABZ97034/C
ID ABZ97034 standard; DNA; 1793 BP.
XX
AC ABZ97034;
XX
DT 17-OCT-2003 (first entry)
XX
DE Human nucleic acid sequence.
XX
KW Human; antisense; lung dysfunction; nasal airway dysfunction;
KW antiinflammatory steroid; ubiqunone; antiinflammatory; antiallergic;
KW antiasthmatic; hypotensive; immunosuppressive; cytostatic; gene therapy;
KW antisense gene therapy; respiratory; lung; adenosine sensitivity;
KW adenosine receptor; bronchodilation; bronchoconstriction; lung allergy;
KW lung inflammation; respiratory disease; ds.
XX
OS Homo sapiens.
XX
FN WO200285308-A2.
XX
PD 31-OCT-2002.
XX
PF 23-APR-2002; 2002WO-US013135.
XX
PR 24-APR-2001; 2001US-0286137P.
XX
PA (EPIG-) EPIGENESIS PHARM INC.

XX NYCE JW, Li Y, Sandrasagra A, Katz E, Pabalan J, Aguilar D;
PI Miller S, Tang L, Shahabuddin S;
XX WPI; 2003-229219/22.
XX Pharmaceutical composition for treating ailments associated with impaired
PT respiration, has oligo(s) antisense to specific gene(s) or its
PT corresponding RNAs, and glucocorticoid or non-glucocorticoid steroid or
PT ubiquinone.
XX Disclosure; SEQ ID NO 12276; 872bp; English.
XX The invention relates to a novel pharmaceutical composition, which has a
CC first active agent comprising an oligonucleotide antisense to the
CC initiation codon, coding region, 5' or 3' end genomic flanking regions,
CC 5' and 3' intron-exon junctions, or regions within 2-10 nucleotides of
CC junctions of genes encoding a polypeptide associated with lung and/or
CC nasal airway dysfunction and a second active agent comprising an
CC antiinflammatory steroid and ubiquinone. A composition of the invention
CC has antiinflammatory, antiasthmatic, hypotensive, immunosuppressive,
CC immunosuppressive, and cytostatic activity. The composition may have a
CC use in antisense gene therapy. The composition is useful for treating or
CC preventing a respiratory, lung or malignant disease or condition, also
CC for enhancing the prophylactic or therapeutic respiratory effect of an
CC antiinflammatory steroid in a subject, for reducing or depleting levels
CC of, or reducing sensitivity to adenosine, reducing levels of adenosine or
CC receptor, producing bronchodilation, increasing levels of ubiquinone or
CC lung surfactant in a subject's tissue, or treating bronchoconstriction,
CC lung inflammation, lung allergies, or a respiratory disease or condition.
CC Note: The sequence data for this patent is not represented in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 1793 BP; 339 A; 546 C; 444 G; 464 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,15e-95 Length: 1793
Score: 1688.00 Matches: 339
Percent Similarity: 99.12% Conservative: 0
Best Local Similarity: 99.12% Mismatches: 3
Query Match: 62.47% Indels: 1
DB: 7 Gaps: 0

US-10-023-523-8 (1-530) x ABZ97034 (1-1793)

QY 189 AnSerGlnLysGlnMetLysLeuGlnLysLysGlnSerGlnLeuValGlnGlnLys 208
DB 1791 AATTACAGAGCATGATGAAGCTCTTACAGAAAAGCAGAGCCAGCTGGTGCAGAGAG 1732
QY 209 AspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeu 228
DB 1731 GACCACCTGCGCGGTGAGCAGACAGAGCC-GTCTGGCCCGCAGCAAGCTTGAGAGCCTA 1673
QY 229 CysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGlyValGlnArgAlaArg 248
DB 1672 TGCCTGTGAGCTGACCGGGCAGCAACCGCTCCCTCAAGGAGAAAGGTGTGACGGGCGCCGG 1613
QY 249 GluGluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIle 268
DB 1612 GAGGAGGAGAGAGGCGCAGGAGGTGACCTCGCCTCCAGTCCAGTGCACATGACATT 1553
QY 269 GlnLeuGlnMetGlnGlnHisAsnGluArgAsnSerLysLeuArgGlnGlnAsnMetGlu 288
DB 1552 CAGCTGCAGATGGAACAGACACAAATGAGCGCAATCCAGCTGCGCCAGAGAAACATGGAG 1493
QY 289 LeuAlaGluArgLeuLysLysLeuLeuGlnValThrGluLeuArgGluGlnHisLeuAsp 308
DB 1492 CTGGCTGAGAGGCTCAAGAGGCTGATGAGCTGATGAGCTGCGGCGAGGAGCATATCGAC 1433
QY 309 LysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAla 328
DB 1432 AAAGTCTTCAAAACACAGAGGACCTTACACAGCAGCTGGTGGATGCCAAGCTCCAGCAGGCC 1373

QY 329 GlnGluMetLeuLysGluAlaGluGlnArgHisGlnArgGluLysAspPheLeuLeuLys 348
DB 1372 CAGGAGATGCTAAAGAGGAGCAGAGAGCGCCACCGAGGAGGATTTTCTCTGAAA 1313
QY 349 GluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlnLysHisLeuLys 368
DB 1312 GAGGAGTAGTATCCAGAGGATGTGTGAGCTGATGAACGACAGAGAGCCACCTGAAG 1253
QY 369 GlnGlnLeuAlaLeuValThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSer 388
DB 1252 CAACAGCTTCCCTATACACAGAGAAAGTTGAGGAGTCCAGAACACACTTCCAAAAGC 1193
QY 389 SerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLys 408
DB 1192 AGCAGAGTATTCCACCATTCAGCAGAGAGATGGAAGAGATGACTAAGAAGATCAAGAAG 1133
QY 409 LeuGluLysGluThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeu 428
DB 1132 CTGGAGAAAAGAACACCATGTACCGGTCCCGTGGGAGCAGCAACAGGCCCTTGCTT 1073
QY 429 GluMetAlaGluGluLysThrValArgAspLysGluLeuGluGluLysLysLysLys 448
DB 1072 GAGATGCTGAGGAGAAAACAGTCCCGGATTAAGAACTGGAGGGCTGCGAGTAAATAATC 1013
QY 449 GlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArg 468
DB 1012 CAACGCTGGAGAGAGCTGTGCGGGCACTGCACAGAGCGCAATGACCTGAACAAGAGG 953
QY 469 ValGlnAspLeuSerAlaGlyGlnGlySerLeuThrAspSerGlyProGluArgArg 488
DB 952 GTACAGACCTGAGTGTGTGCGCCAGGGCTCCCTCCTGACAGAGTGGCCCTGAGAGAGG 893
QY 489 ProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyr 508
DB 892 CCAGAGGGGCTGGGGCTCAAGCACCAGCTCCCCCGGTCACAGAGCGCTTGCTAC 833
QY 509 ProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAla 528
DB 832 CCAGGAGCAGCAGCAGCAGAGATCAGGCCAGATGAGCGGCTCAAGAGCCCTCGGCC 773
QY 529 ArgAla 530
DB 772 AGGGCC 767

RESULT 18
ABZ97033/c
ID ABZ97033 standard; DNA; 1793 BP.
XX AC ABZ97033;
XX DT 17-OCT-2003 (first entry)
XX DE Human nucleic acid sequence.
XX KW Human; antisense; lung dysfunction; nasal airway dysfunction;
KW antiinflammatory steroid; ubiquinone; antiinflammatory; antiasthmatic;
KW antiasthmatic; hypotensive; immunosuppressive; cytostatic; gene therapy;
KW antisense gene therapy; respiratory; lung; adenosine sensitivity;
KW adenosine receptor; bronchodilation; bronchoconstriction; lung allergy;
KW lung inflammation; respiratory disease; ds.
OS Homo sapiens.
XX WO200285308-A2.
XX 31-OCT-2002.
XX 23-APR-2002; 2002WO-US013135.
XX 24-APR-2001; 2001US-0286137P.
XX (EPIG-) EPIGENESIS PHARM INC.

XX Nyce JW, Li Y, Sandrasegra A, Katz E, Pabalan J, Aguilar D;
 PI Miller S, Tang L, Shahabuddin S;
 XX WPI; 2003-229219/22.
 XX Pharmaceutical composition for treating ailments associated with impaired
 PT respiration, has oligo(s) antisense to specific gene(s) or its
 PT corresponding RNAs, and glucocorticoid or non-glucocorticoid steroid or
 PT ubiquinone.
 XX Disclosure; SEQ ID NO 12275; 872pp; English.
 XX The invention relates to a novel pharmaceutical composition, which has a
 CC first active agent comprising an oligonucleotide antisense to the
 CC initiation codon, coding region, 5' or 3' end genomic flanking regions,
 CC 5' and 3' intron-exon junctions, or regions within 2-10 nucleotides of
 CC junctions of genes encoding a polypeptide associated with lung and/or
 CC nasal airway dysfunction and a second active agent comprising an
 CC antiinflammatory steroid and ubiquinone. A composition of the invention
 CC has antiinflammatory, antiasthmatic, hypotensive,
 CC immunosuppressive, and cytostatic activity. The composition may have a
 CC use in antisense gene therapy. The composition is useful for treating or
 CC preventing a respiratory, lung or malignant disease or condition, also
 CC for enhancing the prophylactic or therapeutic respiratory effect of an
 CC antiinflammatory steroid in a subject, for reducing or depleting levels
 CC of, or reducing sensitivity to adenosine, reducing levels of adenosine
 CC receptor, producing bronchodilation, increasing levels of ubiquinone or
 CC lung surfactant in a subject's tissue, or treating bronchoconstriction,
 CC lung inflammation, lung allergies, or a respiratory disease or condition.
 CC Note: The sequence data for this patent is not represented in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1793 BP; 339 A; 546 C; 444 G; 464 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,15e-95 Length: 1793
 Score: 1688.00 Matches: 339
 Percent Similarity: 99.12% Conservative: 0
 Best Local Similarity: 99.12% Mismatches: 3
 Query Match: 62.47% Indels: 1
 DB: 7 Gaps: 0

US-10-023-523-8 (1-530) x AB297033 (1-1793)

QY 189 AsnSerGlnLysGlnMetLysLeuLeuGlnLysGlnSerGlnLeuValGlnGluLys 208
 DB 1791 AATTACAGAGCAGATGAGCTCTCTACGAAAGAGCAGCGCTGTGTCAGAGAG 1732
 QY 209 AspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeu 228
 DB 1731 GACCACCTCGCGGTGAGCAGCAGCAAGCC-GTCCTGGCCCGCAGCAAGCTTGAGAGCCTA 1673
 QY 229 CysArgGluLeuGlnArgHisSerArgSerLysGlnGluGlyValGlnArgAlaArg 248
 DB 1672 TGCCGTGAGCTCAGCGGCACCAACCGCTCCCTCAAGGAAGAGGTGTCAGCGGCCGG 1613
 QY 249 GluGluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIle 268
 DB 1612 GAGGAGGAGGAGCAGCAGCAGGAGGTGACCTTCGCACTTCCAGGTGACACTGAATGACATT 1553
 QY 269 GlnLeuGlnMetGluGlnHisSerGlnLysLeuArgGlnGlnLysMetGlu 288
 DB 1552 CAGCTGTCAGATGGAACAGCACAATGAGCGCAACTCCAAAGCTCGCCCAAGAGCAATGGAG 1493
 QY 289 LeuAlaGluArgLeuLysLysLeuLeuGlnThrGluLeuArgGluGluHisLeuAsp 308
 DB 1492 CTGGCTGAGAGGCTCAGAGAGCTGATTGAGCAGATGATGCTGGCGAGGACATATCGAC 1433
 QY 309 LysValPheLysHisLysAspLeuGlnGlnLeuValAlaLysLeuGlnGlnAla 328
 DB 1432 AAGTCTTCAACACACAGAGACCTACACAGCAGCTGTGGTATGCCAAGCTCCAGCAGGCC 1373

QY 329 GlnGluMetLeuLysGluAlaGluArgHisGlnArgGlnLysAspPheLeuLeuLys 348
 DB 1372 CAGGAGATGCTAAAGGAGGAGCAGAGCGGCACCAGCGGAGAGGATTTTCTCTCTGAAA 1313
 QY 349 GluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLeuLys 368
 DB 1312 GAGCAGTAGAGTCCCAAGAGATGTGTGAGCTGATGAAGCAGCAAGAGACCCACCTGAAG 1253
 QY 369 GlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSer 388
 DB 1252 CAACAGCTTGGCTATACACAGAGAGCTTTGAGAGATTCAGACACACACTTTCNAAGC 1193
 QY 389 SerGluValPheThrThrPheLysGlnGluMetLysMetThrLysLysLysLysLys 408
 DB 1192 AGCGAGGTATTCCACCATTTCAAGCAGAGATGGAAGATGACTAAGAAAGATCAAGAAG 1133
 QY 409 LeuGluLysGluThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeu 428
 DB 1132 CTGGAGAAGAAACCACTATGACCGGTCCCGTGGGAGAGCAGCAACAAAGGCCCTGCTT 1073
 QY 429 GluMetAlaGluLysThrValArgAspLysGluLeuGluGluLeuGlnValLysIle 448
 DB 1072 GAGATGGCTGAGGAGAAAACAGTCCGGGATAAAGAACTGGAGGCGCTGCAGGTAAAAATC 1013
 QY 449 GlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArg 468
 DB 1012 CAACGGCTGGAGAGAGCTGTCCCGGCACTGCAGACAGAGCGCAATGACCTGAACAAAGAG 953
 QY 469 ValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSerGlyProGluArgArg 488
 DB 952 GTACAGGACCTGAGTGTGTCGTCAGGCGCTCCCTCACTCAGAGTGGCCCTGAGAGAGG 893
 QY 489 ProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyr 508
 DB 892 CCAGAGGGCGCTGGGGCTCAAGCACCACCGCTCCCGGAGGTACAGAGCGCGCTTCTTAC 833
 QY 509 ProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAla 528
 DB 832 CCAGAGCACCAGCAGCAGAGAGATCAGGCCAGACTGGGCGCTCAAGAGGCCACCTCCGCC 773
 QY 529 ArgAla 530
 DB 772 AGGCGC 767
 RESULT 19
 ID AAQ66987/c
 XX AAQ66987 standard; cDNA to mRNA; 1854 BP.
 AC AAQ66987;
 XX 25-MAR-2003 (revised)
 DT 19-MAR-1995 (first entry)
 XX
 DE Interleukin 14 gene.
 XX
 KW IL-14; B-cell growth factor; proliferation; differentiation; replication;
 KW non-Hodgkin's lymphoma; systemic lupus erythematosus;
 KW systemic necrotising vasculitis; transplant; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 CDS 73..1770
 FT /*tag= a
 FT
 XX WO9416074-A2.
 XX
 XX 21-JUL-1994.
 PD
 XX
 PF 18-JAN-1994; 94WO-US001101.
 XX
 PR 15-JAN-1993; 93US-00005156.

XX (UYWA-) UNIV WASHINGTON ST LOUIS.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA (TEXA) UNIV TEXAS SYSTEM.
 XX
 PI Ambrus JL, Fauci AS, Ford RJ;
 PI
 XX WPI; 1994-249221/30.
 DR P-PSDB; AAR55800.
 XX
 XX High molecular weight B cell growth factor - able to stimulate B cell
 PT proliferation and inhibit B cell differentiation, useful to treat e.g.
 PT systemic lupus erythematosus.
 XX
 XX Claim 4; Page 75; 95pp; English.
 XX
 CC The sequence is that of a gene encoding a high molecular weight B cell
 CC growth factor which stimulates B cell proliferation and inhibits B cell
 CC differentiation, termed interleukin 14. IL-14 can be used to interfere
 CC with the replication of B cells and to treat B cell non-Hodgkin's
 CC lymphoma, systemic lupus erythematosus and systemic necrotizing
 CC vasculitis. (Updated on 25-MAR-2003 to correct PN field.)
 XX
 XX Sequence 1854 BP; 350 A; 559 C; 456 G; 489 T; 0 U; 0 Other;
 SQ

Alignment Scores:
 Pred. No.: 9,33e-90 Length: 1854
 Score: 1597.00 Matches: 334
 Percent Similarity: 97.66% Conservative: 0
 Best Local Similarity: 97.66% Mismatches: 5
 Query Match: 59.10% Indels: 5
 DB: 2 Gaps: 0

US-10-023-523-8 (1-530) x AAQ66987 (1-1854)

QY 192 LysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGlnLysAspHisLeu 211
 Db 1798 AAGCATGAAGTCTCTACAGAAAGACAGACCCAGCTGGTCAGAGAGACACCTG 1739
 QY 212 ArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGlnSerLysCysArgGlu 231
 Db 1738 CGCGGTGAGCACAGCAAGCC-GTCCTGGCCGCGAGAGGTTGAGAGCCTATGCGGTGAG 1680
 QY 232 LeuGlnArgHisAsnArgSerLeuLysGluGlyValGlnArgAlaArgGlnGluGlu 251
 Db 1679 CTCACGGGCACAAACCGCTCCCTCAAGGAGAGAGGTGTGAGCGGCCCGGAGAGGAG 1620
 QY 252 GluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGln 271
 Db 1619 GAGAGCGCAAGAGGTGACCTCGCACTCCAGGTGACACTGATGACATTGAGCTGCAG 1560
 QY 272 MetGlnGlnHisAsnGluArgAsnSerLysLeuArgGlnGlnAsnMetGluLeuAlaGlu 291
 Db 1559 ATGGAACAGCACAAATGAGCGCAACTCCAAAGCTGCGCCAAAGAACATGAGGTGCTGAG 1500
 QY 292 ArgLeuLysLysLeuLeuGlnGlnTyrGluLeuArgGluGluHisIleAspLysValPhe 311
 Db 1499 AGGCTCAAGAGAGTGTATTGACAGTATGAGCTGCGGAGAGAGCATATCCAAAGTCTTC 1440
 QY 312 LysHisLysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGlnMet 331
 Db 1439 AAACACAAAGGACCTACAAACAGCAGCTGGTGGATGCCAAGCTCCAGCAGGCCAGGAGATG 1380
 QY 332 LeuLysGluAlaGluGlnArgHisGlnArgGluLysAspPheLeuLeuLysGluAlaVal 351
 Db 1379 CTAAAGAGCAGAAAGAGCGGACCAAGCGGAGAGAGATTTCTCTGAAGAGCGAGTT 1320
 QY 352 GluSerGlnArgMetCysGluLeuMetLysGlnGlnGlnThrHisLeuLysGlnGlnLeu 371
 Db 1319 GAGTCCAGAGGATGTGTGAGTGTGATGAGCAGCAGCAAGAGACCCACCTGAAGCAACAGCTT 1260
 QY 372 AlaLeuTyrThrGluLysPheGluGlnPheGlnAsnThrLeuSerLysSerSerGluVal 391

Db 1259 GCCCTATACACAGAGAAAGTTTGAGGAGTCCAGAACACACTTTCCAAAACAGCAGGAGTA 1200
 QY 392 PheThrThrPheLysGlnGlnMetGluLysMetThrLysLysIleLysLysLeuGluLys 411
 Db 1199 TTCACCACTTCAAGCAGAGATGGAAGATGACTAAGAGATCAAGAGAGCTGGACAAA 1140
 QY 412 GluThrThrMetTyrArgSerArgTTPGluSerSerAsnLysAlaLeuLeuGluMetAla 431
 Db 1139 GAAACCACTATGACCGTCCCGTGGGAGAGCAACAAGGCCCTGCTTGAGATGGCT 1080
 QY 432 GluGluLysThrValArgAspLysGluLeuGluGlnGlnValLysIleGlnArgLeu 451
 Db 1079 GAGAGAAAACAGTCCGGATTAAGAACTGAGGGCCTGAGGTAAATAATCCACAGCGCTG 1020
 QY 452 GluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgValGlnAsp 471
 Db 1019 GAGAAGCTGTGCGGGCACTGCACAGAGAGCGCAATGACCTGAACAAGAGGGTACAGGAC 960
 QY 472 LeuSerAlaGlyGlnGlnGlySerLeuThrAspSerGlyProGluArgArgProGlu-Gl 491
 Db 959 CTGAGTCTGTGGCCAGGGCTCCCTCACTGACAGTGGCCCTGAGAGAGGCCAGAGGGG 900
 QY 491 YProGlyVala-GlnAlaProSerSerProArgValThrGluAlaProCysTyrProGlyA 511
 Db 899 CACTGGGGCATCAAGCACCCAGCTCCCCAGGGTTCACAGAAGCGCCTTGCTACCCAGGAG 840
 QY 511 la-ProSerThrGluAlaSerGlyGlnThrGlyProGlnGlnProThrSerAlaArgAla 530
 Db 839 CAACCGAGCACAGAAGCATCAGCCAGACTGG-CCTCAAGAGGCCACCTCGCCAGGGCC 781

RESULT 20
 AAH15042
 ID AAH15042 standard; cDNA; 2045 BP.
 XX
 AC AAH15042;
 DT 26-JUN-2001 (first entry)
 XX
 DE Human cDNA sequence SEQ ID NO:13025.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 OS Homo sapiens.
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-00116126.
 XX
 PR 29-JUL-1999; 98JP-00248036.
 PR 27-AUG-1999; 98JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
 PS Claim 8; SEQ ID NO 13025; 2537pp + Sequence Listing; English.
 XX
 CC The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the

complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention

SQ Sequence 2045 BP; 672 A; 421 C; 506 G; 446 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.96e-70 Length: 2045
Score: 1285.00 Matches:
Percent Similarity: 67.49% Conservative: 71
Best Local Similarity: 53.99% Mismatches: 132
Query Match: 47.56% Indels: 39
DB: 4 Gaps: 8

US-10-023-523-8 (1-530) x AAH15042 (1-2045)

Qy 12 GluGlyAlaGlnArgProSerGlnAlaAlaProAlaValGluAlaGluGlyProGly 31
Db 70 GAGCAGCGCGGAGAGAGCGCGCGCGGAGAGCGGAGAGCGGAGAGCGGCGGAGAG 129
Qy 32 SerSerGlnAlaProArgLys-----ProGluGlyAlaGlnAlaAlaGlnSer 49
Db 130 CGCGAGCGAGCGCGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 189
Qy 50 GlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGluAspLeuSerThr 69
Db 190 GGGCTAGGGGTGAAGCAGATGTTCTGTAATCTCAATCAATCAATCAATCTTCAACAT 249
Qy 70 TyrCysValAspAsnGlnGlyProGlyGluAspGlyAlaGlnGlyGluProAla 89
Db 250 CAA-----GGCTCAAAATTTGTTGGTGGCACAAGTAACAGCAATTCATTGGAAGAGATGA 303
Qy 90 GluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsnGlyGluProGluPro 109
Db 304 GGCAGTGACTTATACAGACAGACAGAGATTTGTTGAGC-----CCA 345
Qy 110 ThrProValValTyrGlyGluLysGluProSerLysGlyAspProAsnThrGluGluLe 129
Db 346 GCATCTGTCAGCAAGAAATCAAGAGAGAGAAATCCCTGGGCGA-----GAAGCT 393
Qy 130 ArgGlnSerAspGluValGlyAspArgAspHisArgArgProGlnGluLysLysAla 149
Db 394 CGAACAGATCCCTGATGTCAGCAGATTCAGAGTGCACAGGAAACAAAGAAAACT 453
Qy 150 LysGlyLeuGlyLysGluLeuThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrPro 169
Db 454 -----TTAGAAAAGAAAGTTTATTACTGATCAAGCCCTCAACACCCCTTCAACCCCA 507
Qy 170 GluGluLysLeuAlaLeuLysLysLysLysLysLysLysLysLysLysLysLysLys 189
Db 508 GAGGAG 567
Qy 190 SerGlnLysGlnMetLysLeuGlnLysLysLysLysLysLysLysLysLysLysLys 209
Db 568 GTTCAGAGAGCAATGAGATCTCTCTAGAAATATGCTGATCTTCTGAGAGAGAGAGAT 627
Qy 210 HisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCys 229

Db 628 CACTTGCAGAGTGAACATAGCAAGCTATCTTGCAGAGCAAGCTAGATCTCTTTGC 687
Qy 230 ArgGluLeuGlnArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGlu 249
Db 688 AGAGAACTTCAGCGCTCAATAAGACGTTAAAGAGGAAATATGCAGCAGGACGAGAG 747
Qy 250 GluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGln 269
Db 748 GAAGNAGAACGACGTAAAGAGCACTGCACATTTCCAGATTACCTTAGATGAATTCAA 807
Qy 270 LeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGlnLysMetGluLeu 289
Db 808 CCCCAGCTGGAGCAGCATGACATCCAAACGCCAACTCCACAGAGAAACATTTGAGCTG 867
Qy 290 AlaGluArgLeuLysLysLeuLysGluGluGluGluGluGluGluGluGluGluGlu 309
Db 868 CGGAGAGAGCTAAAGAGCTCATCGAACGATCCACTGAGGAGAGAGCAATTTGATAG 927
Qy 310 ValPheLysHisLysAspLeuGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGln 329
Db 928 GTGTTCAAACTGAAGAACTGCAACAGCAGCTCGTGATGCCAACTGCAGCAACGACA 987
Qy 330 GluMetLeuLysGluAlaGluGluArgHisGlnArgGluLysAspPheLeuLysGlu 349
Db 988 CAACTGATAAAGAGAGCTGATGAAACATCAGAGAGAGAGAGAGATTTTATTAAAGAA 1047
Qy 350 AlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlnThrHisLysLysGln 369
Db 1048 CGCACAGATCGAGGCACAAATACGAACATGAACAGCAGCAGAGAGTACAACTAAACAG 1107
Qy 370 GlnLeuAlaLeuThrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSer 389
Db 1108 CAGCTTTCTCTTTATGGAATGATGTAATTTCAAGATTTCCAGACTACCATGGCAAAAGCAAT 1167
Qy 390 GluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLysLys 409
Db 1168 GAATGTTTCAACCTTCAGACAGAAATGGAAGAATGACAAAGAAATTAATAAACTG 1227
Qy 410 GluLysGluThrThrMetTyrArgSerArgTrpGluSerSerAsnLysAlaLeuLeuGlu 429
Db 1228 GAAAGAGAAACAATAATTTGGCGTACCATAATGGGAAACAATAATAAAGCACTTCTCAA 1287
Qy 430 MetAlaGluGluLysThrValArgAspLysGluLeuGluGluGluGluGluGluGlu 449
Db 1288 ATGCTGAAG 1347
Qy 450 ArgLeuGluLysLeuLysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgVal 469
Db 1348 CGTITAGAGAGCTGTCAGGCTCTTCAACACAGAAAGAGTGAAGCTCAATGAGAGAGTG 1407
Qy 470 Gln-----AspLeu 472
Db 1408 GAAGTCTGAAAGAGAGAGTATCCATCAAGCGGCCATCAAGCGGCAACAGAGGATTTA 1467
Qy 473 SerAlaGlyGlyGlnGlySerLeuThr-----AspSerGlyProGluArgArgProGlu 490
Db 1468 GCAACACTGTGATGAGCGCTGTACTGCCCTGGATTTCTCAAGAGAGCTGAACACTTCC 1527
Qy 491 GlyPro-----GlyAlaGlnAlaProSer-SerProArgValThrGluAlaProCys 507
Db 1528 TCGAAAAG 1587
Qy 507 sTy-ProGlyAlaPro 512
Db 1588 CAAAAGCCCCGCTCCA 1603

RESULT 21

ABK70301
ID ABK70301 standard; cDNA; 2045 BP.
XX
AC ABK70301;
XX

Db 1108 CAGCTTTCTTTATATGATAGTTTCCAGAAATTCACAGACTACCATGGCAAAAGCAAT 1167
 Qy 390 GluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLysLeu 409
 Db 1168 GAATGTTTCAACCTTCAGACAGAAATGGAAGATGACAAAGAAATTAATAAAACCTG 1227
 Qy 410 GluLysGluThrThrMetTyArgSerArgTTPGILSerSerAsnLysAlaLeuLeuGlu 429
 Db 1228 GAAAAAGAACAAATAATTTGGCGTACCAATGGGAAACATATATAAGCACTTCGCA 1287
 Qy 430 MetAlaGluLysThrValArgAspLysGluLysGluGlyLeuGlnValLysLysGln 449
 Db 1288 ATGCTGAAGAGAAACAGTCCGTGATAAAGAGTCAAGGCCCTTCAAAATAAACTGGAA 1347
 Qy 450 ArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgVal 469
 Db 1348 CGTTAGAGAGCTGTGCAGGGCTCTTCAACAGAAAGGATGAGCTCAATGGAAGGTG 1407
 Qy 470 Gln-----AspLeu 472
 Db 1408 GAAGTCTGAAAGCAGCAGGTATCCATCAAGCGGCCATCAAGCGCGCAACAGGGATTTA 1467
 Qy 473 SerAlaGlyGlyGlnGlySerLeuThr-----AspSerGlyProGluArgProGlu 490
 Db 1468 GCAACCTGTGATGACGCCCTGTACTGCCCTGGATTCTCAACAGGAGCTGAACACTTCC 1527
 Qy 491 GlyPro-----GlyAlaGlnAlaProSer-SerProArgValThr-GluAlaProC 507
 Db 1528 TCGAAAAGACCCCTGGGAGCGCACCTGGAGGTGAGCCCAAGAGTCAAGAGAGCGCTGTG 1587
 Qy 507 sTyProGlyAlaPro 512
 Db 1588 CAAAAGCCCCCGTCCA 1603

RESULT 22

ACD13396
 ID ACD13396 standard; cDNA; 1618 BP.

AC ACD13396;

XX 13-AUG-2003 (first entry)

DE Human DNA encoding a p53 modifier, SEQ ID 67.

XX Human; ss; gene; p53 modifier; cytostatic; cancer; cytostatic;
 KW antiangiogenic; antiapoptotic; p53 pathway; breast cancer; colon cancer;
 KW kidney cancer; lung cancer; ovarian cancer; angiogenesis; cell cycle;
 KW apoptotic disorder; cell proliferation disorder.

XX Homo sapiens.

XX WO200299122-A1.

XX 12-DEC-2002.

XX 03-JUN-2002; 2002WO-US017382.

XX 05-JUN-2001; 2001US-0296076P.

XX 10-OCT-2001; 2001US-0328605P.

XX 15-FEB-2002; 2002US-0357253P.

XX (EXEL-) EXELIXIS INC.

XX Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;

XX WPI: 2003-156859/15.

XX P-PSDB; ABO07222.

XX Identifying modulators of the p53 pathway for use in treating apoptotic
 PT or cell proliferation disorders, comprises screening for agents that
 PT modulate activity of a human ortholog of genes that modify the p53
 PT pathway in Drosophila.

PS Example 2; Page 319-320; 678pp; English.

XX The invention relates to identifying (M1) a candidate p53 pathway
 CC modulating agent, by contacting an assay system comprising a purified HM
 CC polypeptide (human orthologue of genes that modify the p53 pathway in
 CC Drosophila) or nucleic acid with a test agent under conditions where but
 CC for the presence of the test agent, the system provides a reference
 CC activity, and detecting a test agent-biased activity of the assay system.
 CC Also included are modulating (M2) a p53 pathway of a cell (comprising
 CC contacting a cell defective in p53 function with a candidate modulator
 CC that specifically binds to a HM polypeptide comprising an HM amino acid
 CC sequence, where p53 function is restored), modulating (M3) a p53 pathway
 CC in a mammalian cell (comprising contacting the cell with an agent that
 CC specifically binds an HM polypeptide or nucleic acid) and diagnosing (M4)
 CC a disease in a patient (comprising: (a) obtaining a biological sample
 CC from the patient; (b) contacting the sample with a probe for HM
 CC expression; (c) comparing the results with a control; and (d) determining
 CC whether the comparison indicates a likelihood disease). (M1) is useful
 CC for identifying modulators of the p53 pathway. A probe for HM expression
 CC is useful for diagnosing breast, colon, kidney, lung and ovarian cancer,
 CC in a patient where the cancer has greater than 25 % expression level.
 CC Modulators identified by (M1) are useful in a variety of diagnostic and
 CC therapeutic applications, where disease or disorder prognosis is related
 CC to defects in the p53 pathway, such as, angiogenesis, apoptotic or cell
 CC proliferation disorders (e.g. cancer). Another two new methods (M2 and
 CC M3) are useful for modulating the p53 pathway of a cell, thus restoring
 CC the p53 function of the cell, so that the cell undergoes normal
 CC proliferation or progression through the cell cycle. (M2) and (M3) are
 CC also useful for treating defects in the p53 pathway such as angiogenic,
 CC apoptotic or cell proliferation disorders. The present sequence is an HM
 CC nucleic acid encoding a p53 pathway modifying protein
 XX SQ Sequence 1618 BP; 558 A; 330 C; 363 G; 367 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3.08e-70 Length: 1618
 Score: 1280.00 Matches: 257
 Percent Similarity: 80.26% Conservative: 56
 Best Local Similarity: 65.90% Mismatches: 54
 Query Match: 47.37% Indels: 23
 DB: 7 Gaps: 3

US-10-023-523-8 (1-530) x ACD13396 (1-1618)

Qy 146 LysLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThr 165
 Db 9 AGCAACAAGAAAAAACTTTAGGAAAAAGAGTTTATTACTGATGCAAGCCCTAAACACC 68
 Qy 166 LeuSerThrProGluGluLysLeuAlaLeuCysLysLysTyAlaGluLeuGlu 185
 Db 69 CTTTCAACCCAGAGGAGAGCTGGCAGCTCTCTGTAAGAAATATCTGATCTCTGGAG 128
 Qy 186 GluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuVal 205
 Db 129 GAGAGCAGGAGTGTTCAGAAGCAATGAAGATCTCTCAGAGAGAGCAAGCCAGATTGG 188
 Qy 206 GlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeu 225
 Db 189 AAAGAGAAAGTTTCACTTGCAGAGTGAACATAGCAAGGCTATCTTGCAGAGAGCAAGCTA 248
 Qy 226 GluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGlyValGln 245
 Db 249 GAATCTTTTCAGAGAACTTCAGCGTCAATAGACGCTTAAGAGGAGAAATATGCGAG 308
 Qy 246 ArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeu 265
 Db 309 CAGGCAG 368
 Qy 266 AspAspIleGlnLeuGlnMetGlnGlnHisAsnGluArgAsnSerLysLeuArgGlnGlu 285
 Db 369 AATGAAATTCAGCCCGCTGGAGAGCATGATCCACCAAGCCCAATTCAGAGAGAGAA 428
 Qy 286 AsnMetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTyThrGluLeuArgGlu 305

```

Db 429 AACATTGAGCTGGGGAGAGCTAAAGAGCTCATCGAACAGTACCACCTGAGGGGAGAG 488
Qy HisIleAspLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLysLeu 325
Db 489 CACATTGATAGGTGTTCAACAATAGGAAGTCAACAGCAGCTCTGGATGCCAACTG 548
Qy 326 GlnGlnAlaGlnMetLeuLysGluLaGluGluArgHisGlnArgGluLysAspPhe 345
Db 549 CAGCAACACGACACATGATATAAGAGCTGATGAAACATCAGAGAGAGAGAGTTT 608
Qy 346 LeuLysGlnAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlnThr 365
Db 609 TTATTAAAGACGACAGANTCGAGCACAAATAGCAAAATGAAACAGCAAGAGTA 668
Qy 366 HisLeuLysGlnGlnLeuAlaLeuThrPheLysGlnGlnGlnGlnPheGlnAsnThrLeu 385
Db 669 CAACATAAAACAGCAGCTTCTCTTTATATGATAAGTTTGAAGATTCAGACTACCATG 728
Qy 386 SerLysSerSerGluValPheThrPheLysGlnGlnMetCysLysMetThrLysLys 405
Db 729 GCATAAGCAATGACTGTTTCAACCTTCAGACAGAAATGAAAGATGACAAAGAA 788
Qy 406 IleLysLysLeuGluLysGlnThrThrMetTyrArgSerArgTyrGluSerSerAsnLys 425
Db 789 ATTAATAAACTGGAATAAGAAACAATAATTGGCGTACCATAATGGGAAACATAATAA 848
Qy 426 AlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGlnLeuGlnGln 445
Db 849 GCACCTTCTGCAATGCTGGAAGAGAAACAGTCCGTGATAAAGAGTACAAAGCCCTTCAA 908
Qy 446 ValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeu 465
Db 909 ATAAACTGGAACGGTTAGAGAGCTGTGCAGGCTCTTCAGACAGAAAGGATGAGCTC 968
Qy 466 AsnLysArgValGln----- 470
Db 969 AATGAGAGGTGGAGCTCTGAAAGAGCAGGTATCCATCAAGCGGCCCATCAAGCGCG 1028
Qy 471 -----AspLeuSerAlaGlyGlyGlnGlySerLeuThr-----AspSerGlyProGlu 486
Db 1029 AACAGGATTTAGCAACACTGTGTGAGTGCAGCCCTGTACTGCCCTGGATTTCTCACAGGAG 1088
Qy 487 ArgArgProGluGlyPro-----GlyAlaGlnAlaProSer-SerProArgValTh 503
Db 1089 CTGAACACTTCTCGAAAGAGCCCTGGAGGCGCACCTGGAGCTGAGCCCAAGAGTCAG 1148
Qy 503 rGluAlaProCysTyrProGlyAlaPro 512
Db 1149 AGAAGCGCTGTGCAAAAGCCCGCGTCCA 1176
RESULT 23
ACC68992
XX ID ACC68992 standard; cDNA; 4174 BP.
XX AC ACC68992;
XX 09-JUL-2003 (first entry)
XX Human neurotransmission-associated protein NTRAN-14 cDNA SEQ ID NO:39.
XX Human; neurotransmission-associated protein; NTRAN; cytostatic; anti-HIV;
XX antiarteriosclerotic; anticonvulsant; neurotropic; neuroprotective; stroke;
XX cerebroprotective; antiallergic; antiinflammatory; thyromimetic; cancer;
XX antidiabetic; gene therapy; cell proliferative disorder; atherosclerosis;
XX neurological disorder; epilepsy; Huntington's disease; immune disorder;
XX inflammatory disorder; AIDS; allergy; developmental disorder; diabetes;
XX hypothyroidism; Cushing's syndrome; endocrine disorder; infection; gene;
XX ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers

```

```

CDS 309..2363
/*tag= a
/product= "NTRAN-14"
WO2003025129-A2.
27-MAR-2003.
12-SEP-2002; 2002WO-US029219.
14-SEP-2001; 2001US-0322180P.
28-SEP-2001; 2001US-0326096P.
04-OCT-2001; 2001US-0327446P.
26-OCT-2001; 2001US-0345837P.
02-NOV-2001; 2001US-0343903P.
27-NOV-2001; 2001US-0334020P.
07-DEC-2001; 2001US-0340226P.
04-JAN-2002; 2002US-0345008P.
18-MAR-2002; 2002US-0345645P.
10-MAY-2002; 2002US-0379887P.
(INCY-) INCYTE GENOMICS INC.
Honchell CD, Warren BA, Borowsky MB, Griffin JA, Li JX, Lee SY;
Yue H, Forsythe IJ, Marquis JP, Gietzen KJ, Baughn MR, Tran UK;
Lehr-Mason PM, Tang YT, Ramkumar J, Emerling BM, Lee EA, Elliott VS;
Hafalia AJA, Duggan BM, Chawla NK, Kable AE, Chang H, Khare R;
Becha SD, Jin P, Lee S;
WPI; 2003-363137/34.
P-PSDB; ABR43284.
New human neurotransmission associated proteins (NTRAN), useful for
diagnosing, treating and preventing diseases or conditions associated
with the aberrant NTRAN expression e.g. cancer, AIDS, diabetes, epilepsy,
or infections.
Claim 5; Page 230-231; 240pp; English.
ACC68979 to ACC69003 encode the human neurotransmission-associated
proteins given in ABR43271 to ABR43295, designated NTRAN-1 to NTRAN-25
(1). (1) have cytostatic, antiarteriosclerotic, anticonvulsant,
neurotropic, neuroprotective, cerebroprotective, anti-HIV, antiallergic,
antiinflammatory, thyromimetic and antidiabetic activities, and can be
used in gene therapy. The NTRAN polypeptides and polynucleotides are
useful in diagnosing, treating and preventing diseases or conditions
associated with the decreased expression or overexpression of NTRAN, such
as cell proliferative (e.g. cancer, atherosclerosis), neurological (e.g.
epilepsy, Huntington's disease, stroke), immune/inflammatory (e.g. AIDS,
allergies), developmental (e.g. hypothyroidism, Cushing's syndrome) or
endocrine (e.g. diabetes) disorders, or infections. They are also useful
in assessing the effects of exogenous compounds on the expression of
nucleic acid and amino acid sequences of NTRAN. The NTRANs or their
fragments are useful in screening compounds for effectiveness as agonist
or antagonist of the polypeptides, or in altering the expression of the
target polynucleotide and compounds that specifically bind to or modulate
the activity of the polypeptide
Sequence 4174 BP; 1331 A; 924 C; 962 G; 957 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 6,86e-89 Length: 4174
Score: 1265.50 Matches: 276
Percent Similarity: 61.51% Conservative: 90
Best Local Similarity: 46.39% Mismatches: 134
Query Match: 46.84% Indels: 95
DB: 7 Gaps: 10
US-10-023-523-8 (1-530) x ACC68992 (1-4174)
Qy 18 ProSerGlnAlaAlaProAlaValGluAlaGlu----- 28
Db 288 CTTCTGTACTACTCTCCCAAGATGGAGGCTAATCACTCTGAACAGCTCTCAGCGGAACGA 347

```

QY 29 -----GlyProGly---SerSerGlnAlaPro-----ArgLysPro 39
DB 348 CAGTCAACACCTCCAGGTGACAGTTCATCATTAACCCAGTCAACATGGCTCGAGAAGGAA 407
QY 40 GluGlyAlaGlnAlaArgThr-----AlaGlnSerGlyAlaLeuArg 53
DB 408 GATGGCCAGGATCTCCACCCAGTCCACACCCAGTCCACACCCAGAGAGAGAGAGTGCACCCC 467
QY 54 AspValSerGluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysValAsp 73
DB 468 GATATCTCTGACAGCTGAATGACAGCTGGAAGACATCATTAACACTTAT-----GGG 521
QY 74 AsnAsnGlnGlyProGlyGluAspGlyAlaGlnGlyProAlaGluProGluAsp 93
DB 522 TCTGCTGCCAGCACAGCAGGAGAAAGGGCTCTGCCAGGGCCAGTGAGCAGCTGAGAAAT 581
QY 94 AlaGluLysSerArgThrTyrValAlaA:GAsnGlyGluProGluProThrProValVal 113
DB 582 GCAGATCCTCCTGACAC-----GAGGATGGGACTGTGAGGAAACACTGAGAG 632
QY 114 TyrGlyGluLysGluProSerLysGlyAspProAsnThrGluGluIleArgGlnSerAsp 133
DB 633 GCTGGAAGAGAACCCGTGCTTCTGGAGAGCCACCCTCTCAAA-----GAG 680
QY 134 GluValGlyAspArgAspHisArgA:GProGlnGluLysLysAlaLysGlyLeuGly 153
DB 681 CCCGTGAGCAATAGGAGCAAAAA-----TTGGAAGAGAAATCTTAAAGGATTAGGC 734
QY 154 LysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeu 173
DB 735 AAGAAGCCAACTGCTTAATGCAAAATCTGACAAAGTTGCAACACCCGGAAGAAAGTTT 794
QY 174 AlaAlaLeuCysLysLysTyrAlaGluLeuLeuGluGluHisArgAsnSerGlnLysGln 193
DB 795 GATTTTATTATCAAGAGTAGTGTGAATGCTGGATGAACATCGTACTGAGCAAAAGAG 854
QY 194 MetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGly 213
DB 855 TTAAGCTCTCCAAAGAGACAGGTACAAATTCAAAAGAGAAAGGACCAGTTTACAAGT 914
QY 214 GluHisSerLysAlaValLeuAlaArgSerLysLeuLeuSerLeuCysArgGluLeuGln 233
DB 915 GAACACAGCAGAGCTATCTCTCGTCAAGCAAAATTCGAGAGTCTGTGCCGGAGCTGCGAG 974
QY 234 ArgHisAsnArgSerLeuLysGluGlyValGlnAlaArgGluGluGluLys 253
DB 975 AGACAAACAGAGACTCTGAAGGAGAGAGGCGCTTCAGCGGCGAGCTGAGAGAGAGAGAA 1034
QY 254 ArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGlu 273
DB 1035 AGGAAGGAAATCACAAGCCATTTCCAGAGTACCTCAGCGACATCCAGGGGCCAGATCGAG 1094
QY 274 GlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeu 293
DB 1095 CAGCAGAGTGAGCGAAATATGAGCTCTGCAGAGAACACACAGAGCTTGAGAAAGCTG 1154
QY 294 LysLysLeuIleGluGlnTyrGluLeuArgGluGluHisLysAspLysValPheLysHis 313
DB 1155 AAAAGCATCATCGATCATGATGAGTCTCAGAGAGGAGCATCTGACAAAAATATTAAACAC 1214
QY 314 LysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLys 333
DB 1215 AGAAGCTCAGCAGAGAGCTGGTGGATGCAAGAGCTTGAGCGGCCCAAGAAATGATGAAG 1274
QY 334 GluAlaGluGluArgHisGlnArgGluLysAspPheLeuLeuLysGluAlaValGluSer 353
DB 1275 GAACGGAGAGGAGCAGACAAACCGAAGAGGAATATTGCTGAACCGGCGAGCAGAGTGG 1334
QY 354 GlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeu 373
DB 1335 AAACCTTCAGCGGAAAGTCTGAAGGAGCAGAGACAGTCTCTGAGGCTCAGTCTCTC 1394

QY 374 TyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluValPheThr 393
DB 1395 TACTCAGGAAGTTTGAAGAAATTCAGAGCAGCACTACTACTAAAGCAACGAGGTGTTGCC 1454
QY 394 ThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLysLysLysLysGluThr 413
DB 1455 ACGTTCAACAGGAAATGGACAAAACAACTAAGAAAATGAAGAGCTGGAAGAGGACACA 1514
QY 414 ThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeuGluMetAlaGluGlu 433
DB 1515 GCCACATGGAAAGCCCATTTGAGAACTGTAAACAAGACTCTGTGGACATGATTGAAGAG 1574
QY 434 LysThrValArgAspLysGluLeuGluGluGlyLeuGlnValLysLysLysLysLys 453
DB 1575 AAACACTGAGAGCTAAGAAATATGAGTGTCTGTGATGAAATCGGAGGCTAGAGAAC 1634
QY 454 LeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgValGlnAspLeuSer 473
DB 1635 CTCTGCCGTGCTTTTACAAGAGAGAGAGAAACGAACTCCACAAAATAATCAGAGAGCGCAGAA 1694
QY 474 AlaGlyGlnGlySerLeuThrAspSerGlyProGluArgArgProGluGly----- 491
DB 1695 ATATCTGAAAGGATGACCAAAAGTCACCAAACTCCCGATGAGAGCCAGAGTCAAAGTTC 1754
QY 491 -----GAG 491
DB 1755 TCTGTGGATCAAGAGATTGACGAGAGAGGTTAATAGTGTCCAAACCCGCCGTGAAAAAT 1814
QY 491 ----- 491
DB 1815 CTGCCACAGCCTTCATGATAATTCATCATCCAGAGTCAACCCCGCAGTCCAAAGAA 1874
QY 492 -----ProGlyAlaGlnAlaProSer 498
DB 1875 ACCCAACCCGAAATAGCAGTTCTCAGAGAGAGTCTCAGCGCGCTCTCAAGAGAGCCAGAG 1934
QY 499 SerPro-----ArgValThrGluAlaProCysTyrProGlyAlaProSer 513
DB 1935 CAACCCCTCTGATCCCTTCACGGGATTCAGAGAGTCCCTGCTCCCTTAACCTCTCCTCAG 1994
QY 514 ThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAla 528
DB 1995 GCTGAAGCCGAGAGGAGGAGTGTGCTGAACCTCCCTCCCAAGGCC 2039
RESULT 24
AAAX80156
ID AAAX80156 standard; cDNA; 4723 BP.
AC AAAX80156;
XX 16-AUG-1999 (first entry)
DT Neurite extending activity protein encoding cDNA.
DE Neurite extending activity; anti-dementia; memory; brain function;
KW dementia; ds.
XX Mus sp.
OS
FH Key
FT Location/Qualifiers
CDS 196..2226
/*tag= a
JP11147897-A.
XX 02-JUN-1999.
XX 13-NOV-1997; 97JP-00331242.
XX 13-NOV-1997; 97JP-00331242.
XX (AGEN) AGENCY OF IND SCI & TECHNOLOGY.


```

XX 27-OCT-1999 (first entry)
XX
XX cDNA 091-132020 encoding a BRCA1 modulator protein.
XX
XX Modulator protein; BRCA1; tumour suppressor protein; breast cancer;
XX ovarian cancer; cell growth; cell proliferation; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 34..1191
XX /*tag= a
XX /*note= "no termination codon given"
XX
XX US5948643-A.
XX
XX 07-SEP-1999.
XX
XX 13-AUG-1997; 97US-00968751.
XX
XX 13-AUG-1997; 97US-00968751.
XX (ONXX-) ONYX PHARM INC.
XX
XX Rubinfield B, Lingenfelter C, Vuong TT, Polakis PG;
XX
XX WPI; 1999-517952/43.
XX DR
XX P-PSDB; AAY30151.
XX
XX Modulator proteins that bind to and modulate the activity of the BRCA1
XX tumor suppressor gene product, useful for the treatment of ovarian and
XX breast cancer.
XX
XX Claim 1; Fig 3; 35pp; English.
XX
XX The present sequence encodes a modulator protein, that binds to and
XX modulate the activity of the BRCA1 gene product (BRCA1). The BRCA1
XX protein has been characterized as a tumour suppressor protein.
XX Alterations in the amino acid sequence of BRCA1 causes breast and ovarian
XX cancers by removing the controls on cell growth and proliferation.
XX Research has shown that different regions on the BRCA1 molecule have
XX different effects on cell growth and tumour suppression (e.g. full length
XX truncated BRCA1 has no effect on breast cancer cell growth but will
XX inhibit ovarian cancer cell growth). It has been suggested that different
XX host cell factors (e.g. proteins) interact with different regions of the
XX BRCA1 to control its function. The identification of these proteins (e.g.
XX BRCA1MP) will facilitate the development of novel diagnostic methods and
XX new therapeutics for identifying and treating cancers caused by changes
XX in the expression or activity of BRCA1
XX
XX Sequence 1191 BP; 449 A; 228 C; 302 G; 212 T; 0 U; 0 Other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 5,09e-56 Length: 1191
XX Score: 1047.50 Matches: 230
XX Percent Similarity: 69.01% Conservative: 55
XX Best Local Similarity: 55.69% Mismatches: 104
XX Query Match: 38.77% Indels: 24
XX DB: 2 Gaps: 6
XX
XX US-10-023-523-8 (1-530) x AXH86756 (1-1191)
XX
XX QY 3 SerProGlyGlnProGluAlaGlyProGlu-GlyAlaGlnGluArgProSerGlnAlaAl 22
XX
XX Db 12 TCCCCCGGGTGCAGGAATTCCGCACGAGCGCGCGCGAAGAG----- 54
XX
XX QY 22 aProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLys-----ProGl 40
XX
XX Db 55 ----CCGACTAGGCGCGGACCGGCGCGACCGCGACCGCCCGCGCAAGTTTGAAT 110
XX
XX QY 40 uGlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSe 60
XX

```

111 TGGCACAATGGAAGAAGCTGGAAATTTGTGGCTAGCGGTGAAGACAGATATGTTGTGTA 170
60 rArgGlnLeuGluAaspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyProGln 80
CTCTCAATCAAAATGATATCTTTCAACATCAA-----GGCTCAAAATGTGGTGGCACAAG 224
80 yGluAaspGlyAlaGlnGlyGluProAlaGluProGluAaspAlaGluLysSerArgThrTy 100
TAAACAGCAATCATTTGGAAGAGAGATGAAGCGAGTGACTTTATAACAGACAGACAGGAATTT 284
100 rValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGluLysGluProse 120
285 GGTGAGC-----CCAGCATACTGCACGCAAGAANTCAAGACAGGAAAT 326
120 rIysGlyAaspProAsnThrGlnGluIleArgGlnSerAaspGluValGlyAaspArgSpHi 140
327 CCCTGGGGGA-----GAAGCTCGAAGACAGATCCCCCTGATGGTCGACAGCAATTC 374
140 sArgArgProGlnGluLysLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMe 160
375 AGAGTCCACAGCAGCAAGAAAAAACT-----TTAGGAAAGAAGATTTTATTACTGAT 428
160 tGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuLeuCysLysLysTy 180
429 GCAAGCCCTAAACACCTTTCAACCCACAGAGAGAGCTGGCAGCTCTCTGTAAGAAATA 488
180 rAlaGluLeuLeuGluGluHiSargAsnSerGlnLysGlnMetLysLeuLeuGlnLysLy 200
489 TGTGTATCTTCTGAGGAGAGCAGGAGTGTTTCAAGAAGCAATGAAGATCCTGCGAAGAA 548
200 sGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHiSerLysAlaValle 220
549 GCAAGCCCAAGTTGTGAAGAGAAAGATTCACCTGCAGAGTGAACATAGCAAGCTATCTT 608
220 uAlaArgSerLysLeuGluSerLysCysArgGluLeuGlnAcqHisAsnArgSerLeuLy 240
609 GGCAGAAGCAAGCTAGAATCTCTTTGAGAGAACTTCACGCGTCACAAATGAAGACGTTAA 668
240 sGluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHi 260
669 GAGGAAATATGCCAGCGCAGCAGAGGAAGAAGACGCGTATAGAACCAACTGCACA 728
260 sPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSe 280
729 TTTCCAGATTACCTTAATGAAATTCCAAGCCAGCTGGAGCAGCATGACATCCACACGC 788
280 rLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTy 300
789 CAACCTCCACAGGAAAAATTTGAGCTGGGGGAGAGCTTAAGAAGCTCATTCGACAGTA 848
300 rGluLeuArgGluGluHiSirlAspLysValPheLysHisLysAspLeuGlnGlnLe 320
849 CGCACTGAGGGAAGACACATTTGATTAAGTGTTTCAAAACATTAAGGAATGCAACAGCAGCT 908
320 uValAspAlaLysLeuGlnGlnAlaGlnGluMetLysGluAlaGluArgHisGln 340
909 CGTGGATGCCAAACTGCACGAACGACACACATGATAAAGAAGAGCTGATGAATAACATCA 968
340 nArgGluLysAaspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMe 360
969 GAGAGAGAGAGAGTTTTTTATTAAGAAGCGACAGAATCGAGGCACAAATACGAACAAAT 1028
360 tLysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGln 380
1029 GAAACAGCAAGAGTAGTCAACTAAACACAGCAGCTTCTCTTTATATCGATAAGTTTGAAGA 1088
380 uPheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGln 400
1089 ATTCCAGACTACCATGGCAAAAGCAATGAACGTGTTTACACCTTCAGACAGGAAATGGA 1148
400 uLysMetThrLysLysLysLeuGluLysGlu 412
1149 AAAGATGACAAAGAAATTAATAAAAAAAAAAAAAAAAAAAAA 1185

RESULT 27
 ID AAF22469
 AA AAF22469 standard; cDNA; 823 BP.
 AC AAF22469;
 DT 26-MAR-2001 (first entry)
 XX Human breast cancer associated antigen nucleotide sequence SEQ ID NO:48.
 DE Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
 KW cancer associated antigen; cytostatic; cancer vaccine; ss.
 XX Homo sapiens.
 OS
 XX WO200073901-A2.
 FN 07-DEC-2000.
 PD 26-MAY-2000; 2000WO-US014749.
 XX 28-MAY-1999; 99US-0136526P.
 PR 10-SEP-1999; 99US-0153454P.
 XX (LUDW-) LUDWIG INST CANCER RES.
 PA Obata Y;
 PI WPI; 2001-025274/03.
 DR
 XX Nucleic acids encoding breast, gastric and prostate cancer associated
 PT antigen precursors, useful for diagnosing and treating a condition
 PT characterized by expression of an abnormal amount of a protein, e.g.
 PT cancer.
 XX
 PS Claim 50; Page 289-290; 799pp; English.
 XX
 CC AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
 CC represent nucleotide sequences encoding human breast, gastric and
 CC prostate cancer associated antigen precursors (CAAP) respectively.
 CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
 CC represent human breast, gastric and prostate CAAP protein sequence
 CC respectively. CAAPs have cytosolic activity and can be used in the
 CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
 CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
 CC condition characterised by expression of an abnormal amount of a protein,
 CC e.g. cancer
 XX
 SQ Sequence 823 BP; 324 A; 155 C; 192 G; 149 T; 0 U; 3 Other;
 Alignment Scores:
 Pred. No.: Length: 823
 Score: 925.50 Matches: 197
 Percent Similarity: 86.55% Conservative: 41
 Best Local Similarity: 71.64% Mismatches: 37
 Query Match: 34.25% Indels: 5
 DB: Gaps: 0
 US-10-023-523-8 (1-530) x AAF22469 (1-823)
 QY 163 LeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuLysCysLysLysValAlaGlu 182
 DB 2 CTAAACACCTTTCAACCCAGGAGAGAGCTGGCAGCTCTCTGAAGAATATGCTGAT 61
 QY 183 LeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuGluLysLysGlnSer 202
 DB 62 CTTCTGGAGAGAGCAGGAGTGTCTCAGACCAATGAAGATCTCTCAGAGAAAGCAAGCC 121
 QY 203 GlnLeuValGlnLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArg 222
 DB 122 CAGATTGTGAAGAGAAAGTTTCACTTCAGAGATGACATAGCAGGCTATCTTGGCAGA 181

223 SerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGlu 242
 DB 182 AGCAAGCTAGAAATCTTTTCAGAGACTTCAGCGTCACAAATGAAGACGTTTAAGAGAGAA 241
 QY 243 GlyValGlnArgAlaArgGluGluGluGluLysArgLysGluValThrSerHisPheGln 262
 DB 242 AATATGCAGCAGGACGAGAGAGAAAGAACGACGCTAAAGAGCAACTGCACATTTCCAG 301
 QY 263 ValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeu 282
 DB 302 ATTACCTTAATGAATTCAGCCAGCTGGACAGCATGATCCACACAGCCAACTC 361
 QY 283 ArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuLeuGlnLysGluLeu 302
 DB 362 CGACAGGAAACATTCAGCTGGGGGAGAGCTAAAGAGCTCATCGAACAGTACGCACTG 421
 QY 303 ArgGluGluHisLysLeuAspLysValPheLysHisLysAspLeuGlnGlnLeuValAsp 322
 DB 422 AGGAAGAGCAGCATTTGATTAAGGTGTTCAACATGAAGAACTGCAACAGCAGCTCGTGGAT 481
 QY 323 AlaLysLeuGlnGlnAlaGlnMetLeuLysGluAlaGluGluArgHisGlnArgGlu 342
 DB 482 GCCAACTGCAGCAACGACACACTGATAAAGAAAGCTGATGAAACATCATCAGAGAGAG 541
 QY 343 LysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGln 362
 DB 542 AGAGAGTTTATTATAAAGAGCGACAGATCGAGGCACAAATACCAACAAATGAACACAG 601
 QY 363 GlnGluThrHisLeuLysGlnGlnLeuAlaLeuThrGluLysPheGluGluPheGln 382
 DB 602 CAAGAAGTACACTA-AAACAGCAGCTTCTCTTATATGATTAAGTTTGAAGATTNCAG 660
 QY 383 AsnThrLeuSerLysSerSerGluValPheThrPheLysGlnGluMetGluLysMet 402
 DB 661 ACTACCATGTCAGGCAAAAGCAATGAATGTTTCA-ACCTTCAGACAGGAAATGCGAAAGAT 719
 QY 403 ThrLysLysLysLysLysLysLysLysGluThrThrMetTyrArgSerArgTyrGluSer 422
 DB 720 GCCAAGAAATTAATAA-AACTGGGAAAGAAACCACTTAATTTGGGGTTCCTCAATGGGGAAC 778
 QY 423 SerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArg 437
 DB 779 AATATTAAACN-CTTTTGCA-ATGGNTGAAGAAAAAACAGTCCGG 821
 RESULT 28
 AA188232
 ID AA188232 standard; cDNA; 1501 BP.
 XX
 AC AA188232;
 XX
 DT 06-NOV-2001 (first entry)
 DE Human polynucleotide SEQ ID NO 8292.
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation; ss.
 XX Homo sapiens.
 OS
 FN WO200164835-A2.
 PD 07-SEP-2001.
 XX 26-FEB-2001; 2001WO-US004927.
 XX 28-FEB-2000; 2000US-00515126.
 PR 18-MAY-2000; 2000US-00577409.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Drmanac RT;

Qy 421 GluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGlu 440

Db	627	GAGAACTGTAAACAAGACTCTGTGGACATGATTGAAGAGAAGACACTGAGAGCTAAAGAA	686
QY	441	LeuGluGlyLeuGlnValVallysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThr	460
Db	687	TATGAGTGCTTTGTGATGAAATCGGGAGGCTAGAGAACTCTGCGGTGCTTTACAAAGAA	746
QY	461	GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeu	480
Db	747	GAGAGAAACGAACTCCACAAAAAATCAGAGCGCAGAAATATCTGAAAGGATGACCAA	806
QY	481	ThrAspSerGlyProGluArgArgProGluGly-	491
Db	807	AGTCAGCACAACTCCGATGAAGAGCCAGAGTCAAAACGTCTGTGGATCAAGAGATTGAC	866
QY	491	----	491
Db	867	GCAGAGGAGGTATATAGTGTCCAAACCGCGCTGAAAAATCTGCCACAGCCTTCATGATA	926
QY	491	----	491
Db	927	ATTCATCATCCAGAGTCAACCCCGCACCAAGTCCAAAGAAACCCACCGAAATAGGCGAGT	986
QY	492	-----ProGlyAlaGlnAlaProSerSerPro-	500
Db	987	TCTCAGGAGAGTGCTGACGCGCTCTCAAGGAGCCAGAGCAACCCCTCTGATCCCTTCA	1046
QY	501	ArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThr	520
Db	1047	CGGATTTCAGAGAGTCCCTTCCTCCCTAACTCCTCAGGCTGAAGCGGAGGAGCGAGT	1106
QY	521	GlyProGlnGluProThrSerAla 528	
Db	1107	GATGCTGAACCTCCCTCCAGGCG 1130	
RESULT 29			
ADC32449			
ID ADC32449 standard; cDNA; 1501 BP.			
XX	ADC32449;		
XX	18-DEC-2003 (first entry)		
DE	Human novel cDNA contig sequence, SEQ ID NO:2531.		
XX	Human; diagnostic; drug screening; forensics; gene mapping;		
KW	biodiversity assessment; Parkinson's disease; Alzheimer's disease;		
KW	neurodegenerative diseases; anaemia; platelet disorder; wound; burns;		
KW	ulcers; osteoporosis; autoimmune disease; cancer;		
KW	molecular weight marker; food supplement; antiparkinsonian; nootropic;		
KW	neuroprotective; antianemic; anticoagulant; thrombolytic; vulnerary;		
KW	antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;		
KW	gene therapy; chromosome 6; ss.		
XX	Homo sapiens.		
OS	WO2003029271-A2.		
PN	10-APR-2003.		
PD	24-SEP-2002; 2002WO-US030474.		
PF	24-SEP-2001; 2001US-0324631P.		
XX	(HYSE-) HYSEQ INC.		
XX	Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;		
PI	Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;		
PI	Halley-Vicente D, Drmanac RT;		
XX	WPI; 2003-371981/35.		
DR	P-PSDB; ADC32449.		
XX	New polynucleotide and polypeptide useful for diagnosing, preventing or		

XX New polynucleotide and polypeptide useful for diagnosing, preventing or

XX DE Human low density lipoprotein binding protein 3 (LBP-3) gene.
 XX DE
 KW KW Low density lipoprotein binding protein 3; LBP-3; LDL; human;
 KW KW atherosclerosis; antiarteriosclerotic; gene therapy; diagnosis; vaccine;
 KW ds.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 XX CDS 2100..16332
 FT FT /*tag= a
 FT FT /note= "contains introns"
 FT FT 2100..2268
 FT FT /*tag= b
 FT FT 2269..2996
 FT FT /*tag= c
 FT FT 2997..3332
 FT FT /*tag= d
 FT FT 3333..6309
 FT FT /*tag= e
 FT FT 6310..6401
 FT FT /*tag= f
 FT FT 6402..9738
 FT FT /*tag= g
 FT FT 9739..9909
 FT FT /*tag= h
 FT FT 9910..11839
 FT FT /*tag= i
 FT FT 11840..12034
 FT FT /*tag= j
 FT FT 12035..14095
 FT FT /*tag= k
 FT FT 14096..14215
 FT FT /*tag= l
 FT FT 14216..14443
 FT FT /*tag= m
 FT FT 14444..14518
 FT FT /*tag= n
 FT FT 14519..14944
 FT FT /*tag= o
 FT FT 14945..15037
 FT FT /*tag= p
 FT FT 15038..15801
 FT FT /*tag= q
 FT FT 15802..15897
 FT FT /*tag= r
 FT FT 15898..16638
 FT FT /*tag= s
 FT FT 16639..16932
 FT FT /*tag= t
 XX WN 200164874-A2.
 XX PD 07-SEP-2001.
 XX PF 28-FEB-2001; 2001WO-US006356.
 XX PR 02-MAR-2000; 2000US-00517849.
 XX PR 14-JUL-2000; 2000US-00616289.
 XX PA (BOST-) BOSTON HEART FOUND INC.
 XX PI Lees AM, Lees RS, Law SW, Arjona AA;
 XX DR WPI: 2001-565505/63.
 XX DR P-PSDB; AAB2809.
 XX PT New isolated low density lipoprotein binding polypeptide for treating,
 XX PT diagnosing and/or identifying therapeutic agents for atherosclerosis.
 XX PS Example 4; Fig 24; 143pp; English.
 XX XX

CC The present sequence is that of genomic DNA encoding novel human low
 CC density lipoprotein binding protein 3 (LBP-3, see AAB2809). The DNA was
 CC isolated from a human genomic library by screening with LBP-3 CDNA (see
 CC AAB2809). The open reading frame spans 10 exons. Human LBP-2 nucleic
 CC acids are among claimed polynucleotides of the invention that encode
 CC novel polypeptides, termed LBPs, capable of binding to native and
 CC methylated LDL. Also claimed are isolated LBP polypeptides, and
 CC biologically active fragments and analogues of them, as well as
 CC expression vectors, cells and methods of producing the LBPs. Methods of
 CC determining if an animal is at risk for atherosclerosis, methods for
 CC evaluating an agent for use in treating atherosclerosis, and methods for
 CC treating a cell having an abnormality in structure or metabolism of LBP
 CC are claimed. Pharmaceutical compositions comprising an LBP polypeptide or
 CC nucleic acid, and vaccine compositions, are also claimed
 XX
 SQ Sequence 22255 BP; 5195 A; 5302 C; 5924 G; 5834 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.74e-34	Length:	22255
Score:	721.50	Matches:	224
Percent Similarity:	23.78%	Conservative:	1
Best Local Similarity:	23.68%	Mismatches:	1
Query Match:	26.70%	Indels:	721
DB:	5	Gaps:	4

US-10-023-523-8 (1-530) x AAB26497 (1-22255)

QY	305	GluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeuValAspAlaLys	324
DB	14093	CAGCATATCGACAAAGTCTTCAACACACAGACCTACACAGCAGCTGGTGGTCAAG	14152
QY	325	LeuGlnGlnAlaGlnGlnMetLeuLysGluAlaGluArgHisGlnArgGluLysAsp	344
DB	14153	CTCCAGCAGGCCCGCAGGAGATGCTAAAGAGGACAGAGCGGCCACCGCGGAGAGGAT	14212
QY	345	Phe-----	345
DB	14213	TTTGTAGGCTCAGGCCCGCAGGTTGGGGTGTGGAGGAGACAGCTGGGCTCT	14272
QY	345	-----	345
DB	14273	GGCTCAGCTCATACCGGGTTATATGGGAGAAGTCTGGCCAGACAGGACAGATTCTT	14332
QY	345	-----	345
DB	14333	GAGTACCAGTCTGAGACAGCAGAGCCCTCAGTGGGTCTGGTGTGGCTAAACACAAA	14392
QY	346	-----	LeuLeuLys 348
DB	14393	CATAGCCCTGGGGCTTCTGACAGGATCTGGGTTCTGTCTTGGAAATAGCTCCTGAAA	14452
QY	349	GluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLeuLys	368
DB	14453	GAGCAGTAGATCCACAGAGATGTGTAGCTGTATGAGCAGCAGAGACCCACCTCAAG	14512
QY	369	Gln-----	369
DB	14513	CAACAGGTGAGACATATAACCTGACCCCTGTGCTTCAAGTTTCCCTCACTGGGCCCAT	14572
QY	369	-----	369
DB	14573	CTGGGGGTAGTGAATGGGACCCCTCATTTAGGACTGGGTGTGTCTGTGCTGTATGAC	14632
QY	369	-----	369
DB	14633	GCCTTGGTTGAGCTTAGTGGGCTCAGAGGACTTTCATTTGTAGCTCAGAAATGATTGCT	14692
QY	369	-----	369
DB	14693	TTTGAGGAGTAGGACACAGAGAGTTTGAATAATCAACATAAAGGCAAAATAAAGTCAAC	14752
QY	369	-----	369

AAV88391	AAV88391 standard; cDNA; 529 BP.
XX AC	AAV88391;
XX DT	12-FEB-1999 (first entry)
XX DE	EST clone GP232.
XX KW	Expressed sequence tag; secreted protein; haematopoiesis regulator;
XX KW	tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
XX KW	chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
XX KW	receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
XX OS	Homo sapiens.
XX EN	W09845437-A2.
XX PD	15-OCT-1998.
XX PF	10-APR-1998; 98WO-US006956.
XX PR	10-APR-1997; 97US-00837312.
XX PA	(GEMY) GENETICS INST INC.
XX PI	Jacobs K, McCoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;
XX PI	Spaulding V, Agostino MJ;
XX DR	WPI; 1999-070078/06.
XX PT	New polynucleotides encoding human secreted proteins - derived from e.g.
XX PT	human blood, kidney, foetal lung, placenta, testes, brain, ovary,
XX PT	pituitary, retina and colon cDNA libraries.
XX PS	Claim 1; Page 381; 641pp; English.
XX CC	The present sequence represents an expressed sequence tag (EST), and is a
XX CC	polynucleotide of the invention. The polynucleotides of the invention are
XX CC	all secreted EST sequences isolated from a variety of human tissue
XX CC	sources. The EST sequences and proteins encoded by them are predicted to
XX CC	have useful biological activities which would make them suitable for
XX CC	treating, preventing or ameliorating medical conditions in humans and
XX CC	animals, although no supporting data is given. Suggested activities
XX CC	include nutritional activity, immune stimulating or suppressing activity,
XX CC	haematopoiesis regulating activity, tissue growth activity,
XX CC	activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX CC	and thrombolytic activity, receptor/ligand activity, anti-inflammatory
XX CC	activity, cadherin/tumour invasion suppressor activity, tumour inhibition
XX CC	activity. The EST sequences are also stated to be useful for gene therapy
XX SQ	Sequence 529 BP; 151 A; 146 C; 168 G; 64 T; 0 U; 0 Other;

Alignment Scores:	
Pred. No.:	9.69e-36
Score:	713.00
Percent Similarity:	99.28%
Best Local Similarity:	98.56%
Query Match:	26.33%
DB:	2
	Gaps:
	0

US-10-023-523-8 (1-530) x AAV88391 (1-529)

QY	1	LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln	20
Db	104	AAAGCAGCCAGGACACACCGAAGCAGGAGCCAGGAGCGGCCACCCAG	163
QY	21	AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu	40
Db	164	CGCGGCTCTCGAGTAGAAGCAGAGAGGTCCCGGACGACGCAGGCTCTCTGGAGCCGGAG	223
QY	41	GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer	60

224 GGGGCTCAAGCCAGAACGGCTCAGTCTCGGGCCCTTCTGTGATGTCTGTGAGGAGCTGAGC 283

61 AtgGlnLeuGluAspIleuSerThrTyrCysValaspAsnAsnGlnGlyGlyProGly 80

284 CGCAACTCGAAGACATCTAGCAGCACACTGTGTGGACAATAACCAAGGGGGCCCCGGC 343

81 GluAspGlyValaGlnGlyGluProAlaGluProGluAspAlaGluLysSerA:gthrTyr 100

344 GAGGATGGGGCACAGGGTGAGCGGCTGAACCCGAAGATGCAGAGAAGTCCCGGACCTAT 403

101 ValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGluLysGluProSer 120

404 GTGGCAAGGAATGGGGAGCCTGAACCACTCCAGTAGTCAATCGAGAGAAGAACCTCC 463

121 LysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspA:Gasp 139

464 AAGGGGGATCCAAACACAGAGAGATCCGGCAGAGTGCAGAGGTGGAGACCGAGAA 520

RESULT 33

ACD13448

ID ACD13448 standard; DNA; 115756 BP.

XX AC ACD13448;

XX AC ACD13448;

XX 13-AUG-2003 (first entry)

XX Human DNA encoding a p53 modifier, SEQ ID 68.

DE Human; ds; gene; p53 modifier; cytostatic; cancer; cytostatic;

XX Human; ds; gene; p53 modifier; cytostatic; cancer; cytostatic;

KW antiangiogenic; antiapoptotic; p53 pathway; breast cancer; colon cancer;

KW kidney cancer; lung cancer; ovarian cancer; angiogenesis; cell cycle;

KW apoptotic disorder; cell proliferation disorder.

XX

OS Homo sapiens.

OS

XX WO200299122-A1.

PN

XX 12-DEC-2002.

PD

XX

XX 03-JUN-2002; 2002WO-US017382.

PF

XX

XX 05-JUN-2001; 2001US-0296076P.

PR

XX 10-OCT-2001; 2001US-0328605P.

PR

XX 15-FEB-2002; 2002US-0357253P.

PR

XX (EXEL-) EXELIXIS INC.

PA

XX

XX Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;

PI

XX WPI; 2003-156859/15.

PT

XX P-PSDB; ABO07223.

DR

XX

XX Identifying modulators of the p53 pathway for use in treating apoptotic

PT or cell proliferation disorders, comprises screening for agents that

PT modulate activity of a human ortholog of genes that modify the p53

PT pathway in Drosophila.

XX

XX Example 2; Page 320-351; 678pp; English.

XX

XX The invention relates to identifying (M1) a candidate p53 pathway

CC modulating agent, by contacting an assay system comprising a purified HM

CC polypeptide (human orthologue of genes that modify the p53 pathway in

CC Drosophila) or nucleic acid with a test agent under conditions, where but

CC for the presence of the test agent, the system provides a reference

CC activity, and detecting a test agent-biased activity of the assay system.

CC Also included are modulating (M2) a p53 pathway of a cell (comprising

CC contacting a cell defective in p53 function with a candidate modulator

CC that specifically binds to a HM polypeptide comprising an HM amino acid

CC sequence, where p53 function is restored), modulating (M3) a p53 pathway

CC in a mammalian cell (comprising contacting the cell with an agent that

CC specifically binds an HM polypeptide or nucleic acid) and diagnosing (M4)

CC a disease in a patient (comprising: (a) obtaining a biological sample

CC

from the patient; (b) contacting the sample with a probe for HM expression; (c) comparing the results with a control; and (d) determining whether the comparison indicates a likelihood disease). (M1) is useful for identifying modulators of the p53 pathway. A probe for HM expression is useful for diagnosing breast, colon, kidney, lung and ovarian cancer, in a patient, where the cancer has greater than 25 expression level. Modulators identified by (M1) are useful in a variety of diagnostic and therapeutic applications, where disease or disorder prognosis is related to defects in the p53 pathway, such as, angiogenesis, apoptotic or cell proliferation disorders (e.g. cancer). Another two new methods (M2 and M3) are useful for modulating the p53 pathway of a cell, thus restoring the p53 function of the cell, so that the cell undergoes normal proliferation or progression through the cell cycle. (M2) and (M3) are also useful for treating defects in the p53 pathway such as angiogenic, apoptotic or cell proliferation disorders. The present sequence is an HM nucleic acid encoding a p53 pathway modifying protein

[illegible]

305	Qy	GluHisIILAspLysValPheIysHisLysAspLeuGlnGlnGlnLeuValAspAlaLys	324
81636	Db	CGCATATCGACAAAGCTCTCAACACAAAGGACCTACACAGCAGCTGGTGATGCCAAG	81695
325	Qy	LeuGlnGlnAlaGlnGlnMetLeuLysGlnAlaGluGluAArgHisGlnArgGluLysAsp	344
81696	Db	CTCCAGCAGCCCGAGGAGATGCTTAAGAGAGGACAGAGACCGCACACACGGGGAGAGAGAT	81755
345	Qy	Phe	345
81756	Db	TTTGTGAGGCTCAGGCCCCAGGGTTGGGTGGGGTGGGAGAGACAGGCTGGGCTCTGG	81815
345	Qy		345
81816	Db	CTCAGCTCATAGCCGGTTATATGGGAGAAAGTCTGCCACAGACCAGGCACAGATTCCCTGA	81875
345	Qy		345
81876	Db	GTACCAGTCTGAGACAGGAAGCCTCAGTGGGTCTGCTTGTGGCTAAAAACCAACA	81935
346	Qy		346
81936	Db	TAGCCCTGGGGCTTCTCACAGGATCTGGGGTTCTGTCTGGAAATAGCTCTCTGAAGA	81995
349	Qy	uLaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLeuLysGln	369
81996	Db	GGCAGTAGAGTCCCAGAGATGTGTGAGCTGATGAAGCAGCAGAGACCCACCTCAAGCA	82055
369	Qy	nGlnLeu	371
82056	Db	ACAGGT-GAGAGCATATAACCTGACCCCTGTGCCTTCAAGTTTCCCTCACTGGGCCCCATC	82114
371	Qy		371
82115	Db	CTGGGGTAGTGAAATGGGACCTCATTTAGGACTGGCTGTGCTCTGGCTGCTATGACG	82174
371	Qy		371
82175	Db	CCTTGGTTGAGCTTTGTTCTCTCCGACCTGCACAGTACCTATGTTGGTGGTGACAGGT	82234
371	Qy		371
82235	Db	AGTTAGGTGGGCTCAGAGGACTTCATTTTGTAGCTCAGAAATGTTATTTGCTTTGAGGAGGT	82294

```
Db 83374 TGCTGTGCATTAACCTAGATGACTAAGAGATCAAGAAGCTGGAGAAAGAACCCACCAT 83433
QY 415 tTyArgSerArgTrpGluSerSerAsnLysAlaLeuLeuGluMetAlaGluGlu----- 433
Db 83434 GTACCGGTCCCGTGGGAGAGACGACCAAGGCCCTGCTTGAGATGGCTGAGAGAGTGGG 83493
QY 433 ----- 433
Db 83494 CTGCTGTGATGTCGAGCCAGGCGTGGGGTGTGCACTTAGCGCATATCAGGCGCTTTCCT 83553
QY 433 ----- 433
Db 83554 GTATGTTCTACCCATCAGTGACACAGCTAGCATGAGGTAGAGGTGAGATTCCACACAAAT 83613
QY 433 ----- 433
Db 83614 GTCCAAAGTCCAAAGTTAATGCTGTTCTCTCCCATGGGAGGTGGTGAGCCAGTGGTAGG 83673
QY 433 ----- 433
Db 83674 TCTCCAGTGGAGTGAAGGAGCAAAATGGAAGAAAGGAATAAAGAGCAGAAAAAACGG 83733
QY 433 ----- 433
Db 83734 GTGCCAGTGTGCTGCTGTTTACATGTAAGCAGCCAGGTAGTGTGATTTCCACAGC 83793
QY 433 ----- 433
Db 83794 TTGTAATGAGAAGAAAGAACTAAGATGGAGCACTCAAGCAGACCTTGCTGA 83853
QY 433 ----- 433
Db 83854 AAGTTTGGGTTTTTTTGTCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 83913
QY 433 ----- 433
Db 83914 GAACCTTCTCTTGTAGCTCTGAGGATGATATTAGTAGTCTGCTGTTTATAGATGAGACAGGC 83973
QY 433 ----- 433
Db 83974 TCAAAAGTCAAGTCTCTTTGCCAAGTCACTGCTGATGATTAATGAGGAATACGTTATCTC 84033
QY 433 ----- 433
Db 84034 CAAGCGTGCCCTTTTCTGCACCATGCTGCCACCTGACAGCCTAGTATGCTTCAA 84093
QY 433 ----- 433
Db 84094 CTAGGACTGTTTCTTAGAGGGGGCCAGCTTTGGACTCGTCTCTCAGCCTTTGTTAA 84153
QY 433 ----- 433
Db 84154 GTGTTTGGCCCAAGTGTGATGTTAAGTGGGAGGTTGATGGGCGACCGCACTGAAGTC 84213
QY 434 -----LysThrValArgAspLysGluLeuGluGlyLeuGlnValIysII 448
Db 84214 TCATTTCTTCCCTAGAAAAACAGTCGGGATTAAGAACTGGAGGGCTGCGAGGTAAATAAT 84273
QY 448 eGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysAr 468
Db 84274 CCACGGCTGGAGAGCTGTCCGGGCACTGCACACAGAGCGCAATGACCTGACACAGAG 84333
QY 468 gValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSerGlyProGluArgAr 488
Db 84334 GGTACAGACCTGAGTGTGCTGTGGCCAGGGCTCCCTCACTGACAGTGGCCCTGAGAGGAG 84393
QY 488 gProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTy 508
Db 84394 GCCAGAGGGGCTGGGCTCTCAGGACCCAGCTCCCGAGGGTCAACAAGCGCTTGCTA 84453
QY 508 rProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAl 528
```

```
Db 84454 CCCAGGACCCGACGACAGACAGCATCAGGCCAGCTGGGCTCAAGAGCCACCTCCGC 84513
QY 528 aArgAla 530
Db 84514 CAGGCCC 84520
RESULT 34
AAS66673
ID AAS66673 standard; cDNA; 1233 BP.
XX
AC AAS66673;
XX
DI 13-FEB-2002 (first entry)
XX
DNA encoding novel human diagnostic protein #2477.
Human; chromosome mapping; gene mapping; gene therapy; forensic;
food supplement; medical imaging; diagnostic; genetic disorder; ss.
Homo sapiens.
XX
WO200175067-A2.
XX
11-OCT-2001.
XX
30-MAR-2001; 2001WO-US008631.
XX
31-MAR-2000; 2000US-00540217.
XX
23-AUG-2000; 2000US-00649167.
XX
(HVSE-) HYSEQ INC.
XX
Dmanac RT, Liu C, Tang YT;
XX
WPI; 2001-639362/73.
XX
P-PSDE; ABG02486.
XX
New isolated polynucleotide and encoded polypeptides, useful in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits and to assess
biodiversity.
Claim 1; SEQ ID NO 2477; 103pp; English.
The invention relates to isolated polynucleotide (I) and polypeptide (II)
sequences. (I) is useful as hybridisation probes, polymerase chain
reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
and in recombinant production of (II). The polynucleotides are also used
in diagnostics as expressed sequence tags for identifying expressed
genes. (I) is useful in gene therapy techniques to restore normal
activity of (II) or to treat disease states involving (II). (II) is
useful for generating antibodies against it, detecting or quantitating a
polypeptide in tissue, as molecular weight markers and as a food
supplement. (II) and its binding partners are useful in medical imaging
of sites expressing (II). (I) and (II) are useful for treating disorders
involving aberrant protein expression or biological activity. The
polypeptide and polynucleotide sequences have applications in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits to assess biodiversity
and to produce other types of data and products dependent on DNA and
amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
coding sequences of the invention. Note: The sequence data for this
patent did not appear in the printed specification, but was obtained in
electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences
XX
Sequence 1233 BP; 438 A; 230 C; 287 G; 278 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 5.55e-31 Length: 1233
Score: 642.50 Matches: 153
Percent Similarity: 55.61% Conservative: 55
Best Local Similarity: 40.91% Mismatches: 85
```

Query Match:	23.78%	Indels:	81
DB:	5	Gaps:	11
US-10-023-523-8 (1-530) x AAS666673 (1-1233)			
QY	167	SerThrProGluGluLysLeuAlaLeuCys-----LysLysTyrAlaGluLeuLeu	184
DB	253	AGCAAAATGGAAGAA-----GCTGGACTTTCTGGTTAAGAGAGAAAGCAGATATGTTG	306
QY	185	-----GluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysGln	201
DB	307	TGTAACCTCTGAATCAT-----GATATCTTCAACATCAAGAC	345
QY	202	Ser-----GlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaVal	219
DB	346	TCAAAATTCAGTCCCAACAGTAATAAATATTTATTCGAAGATGAAGAGGCCGCTGACTTT	405
QY	220	LeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeu	239
DB	406	ATAACAAAG-----AACAGAGTTGG	426
QY	240	LysGluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSer	259
DB	427	---GAGGAACATATGACGCAAGAAAAGAGAGGAGGAAGTCTTTAAAGAAAGTAACGTCA	483
QY	260	HisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsn	279
DB	484	CATTTCCAATTAACCTTAACCTGAACCTCAAGCCAGCTGGAACTGAATGAATACAAAT	543
QY	280	SerLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuLysGluGln	299
DB	544	GCCAACTGCAGCAGAGAAACATGGAATGGGAGAAAGCTTAAGAAAGCTCACTGACCAG	603
QY	300	TyrGluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGln	319
DB	604	TATGCATCAGGGAAGAGCAAAATTAATAAAGCGTTCAAAACATTAAGAAATTAACGACAA	663
QY	320	LeuValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHis	339
DB	664	CTTGATAGTCCAGACTTCAACAAACAGCAGCTGATAAAGAAAGCTGATGAAGACAT	723
QY	340	GlnArgGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeu	359
DB	724	CAGAGAGAGAGAGAGTTTATTAATAAAGAAAGCAAGAAATCGAGGCACAAATATGAAGAA	783
QY	360	MetLysGlnGluThrHisLysLysGlnGlnLeuAlaLeuTyrThrGluLysPheGlu	379
DB	784	ATGAACAAAGAGAGAGCAACAACTGAAGAGAGAGCTTTTCCTTTATGATGATGAATGAA	843
QY	380	GluPheGlnAsnThrLeuSerLysSerGluValPheThrThrPheLysGlnGluMet	399
DB	844	GAATTCAGACTACCATGCAAAACCAATGAACCTTTTACAGCTTCAAGCAGGAAACG	903
QY	400	GluLysMetThrLysLysLysLysLysLeuGluLysGluThrMetTyrArgSerArg	419
DB	904	GAAGAAG-----	909
QY	420	TrpGluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLys	439
DB	910	-----AAACTTCTCGTGATAA	927
QY	440	GluLeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGln	459
DB	928	AATTATAAGTCTTTCAAAATAAATGGAGCGTTAGAGAGCTGTACAGGCTCTTCAA	987
QY	460	ThrGluArgAsnAspLeuAsnLysArgVal-----	469
DB	988	ATGAAGAGGAATGAATCTAGTGAAGAACTGGGAATTCGAAGGCGAGCTCTGTGAA	1047
QY	470	-----GlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSerGlyProGlu	486
DB	1048	GTAGCAGATGATAGATTTAGCAGTGCTGTGACCGCATCTCTGTGCTGATCTCTCC	1107

Alignment Scores:
Pred. No.:
Score:

db 54 GAGAAGCTAACCAAGCGCAAAAAAGTGGCGCGGAGGAGCAAGCAGCGGCCAGAGATTG 113

```
QY 147 LysLysAlaLysGlyLeuGlyLysGluIleThrLeuMetGlnThrLeuAsnThrLeu 166
DB 114 GAGAG-----TTGGTCATGAATCCCTTCACGAATGC 146
QY 167 SerThrProGluGluLysLeuAlaLeuCysLysLysTyraLalaGluLeuGluGlu 186
DB 147 CCTCGCGCAGGAGAGAGTGAACCT-----CTGCTGCAACGC 185
QY 187 HisArgAsnSerGlnLys-----GlnMetLysLeuLeuGlnLys 199
DB 186 CATGTGCAGCAGCAAAAGAAATGTAAGCCGGCTGACGGCGAACTCGGTGTCCTGAGCGC 245
QY 200 LysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaVal 219
DB 246 CAATGCGAGTCAGCAGAGGAGAGAGAGAGTTCACGGGATCTTAACAGAGCGTC 305
QY 220 LeuAlaArgSerLysLeuGluSerLysCysArgGluLeuGlnArgHisAsnArgSerLeu 239
DB 306 CTGATCGGAGCAAGCTGCGAGAGGTGTCGCCGCAACAGCAGCGCATATAAGTCCGTA 365
QY 240 LysGluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSer 259
DB 366 AGAAGCAGAGCTGCTGCAATCAATGAGTGGAGGAGCGCGCAAGAGAGCGCAGACA 425
QY 260 HisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsn 279
DB 426 AAGTTCCTCAAGCTCCCTGAACCATGTCCTGCAAGAGTGCCTGCCCAAGCAACAGCAGAGAAC 485
QY 280 SerLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGluGln 299
DB 486 ATCAAGCTCGCGGATTACCAATCGAGATGACCAAGAGTTAAAGCTGCTGCCCGCAGAG 545
QY 300 TyrGluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGln 319
DB 546 TACCAACAGGAGGAGCAGCCTGAGAGAGTCAATGACAGGTCAGTTGGAGCCCGAG 605
QY 320 LeuValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluGluHis 339
DB 606 CTGCAACAGCAGCAAGCTGCAAAATGCCAGGTGGAGGCTGCCATGCAAAAGAGGATATTG 665
QY 340 GlnArgGluLysAspPheLeuLysLysGluAlaValGluSerGlnArgMetCysGluLeu 359
DB 666 AGCAAGAAACACAGATGCTGCTGAGAACTGTATGCAAGCAACGTCGCCATAAGGAT 725
QY 360 MetLysGlnGlnGluThrHisLysLysGlnGlnLeuAlaLeuTyThrLysPheGlu 379
DB 726 CTCCAGATCGGAGCAGCAACCACTGAAGGAGCACTGAACATATACACGCCCAAGTACCAT 785
QY 380 GluPheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMet 399
DB 786 GACTTCCAGAGTCGCTGAGAGAGTCCAAATGAGGTTTTCGGCAGCTACAAAGTGGAGCTT 845
QY 400 GluLysMetThrLysLysIleLysLysLysLeuGluLysGluThrThrMetTyArgSerArg 419
DB 846 GAGAAGATGTCACAGCACACTAAGAAGATCGAGAGGAGCGCTCGGCTGGCGGCAAGAG 905
QY 420 TrpGluSerSerAsnLysAlaLeuGluMetAlaGluGluLysThrValArgAspLys 439
DB 906 TACGAGAAAGCTAATGCCATGCTGATGCTGATTTAGCCACAGAAAGAGTCTGACGAGCG 965
QY 440 GluLeuGluGluValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGln 459
DB 966 CACTCAGAGCGCTGAGAGAGCAATACAGCAGCTGCGAGAGCTACTCGGTGCTCTACAG 1025
QY 460 ThrGluArgAsnAspLeuAsnLysArgValGlnAsp 471
DB 1026 CTGGAGAGAGCAACACTGACCAAGTGCCTCGCGGAT 1061
```

RESULT 37

AAS66674

ID AAS66674 standard; cDNA; 2526 BP.

XX

```
AC AAS66674;
XX 13-FEB-2002 (first entry)
XX DNA encoding novel human diagnostic protein #2478.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG02487.
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX Claim 1; SEQ ID NO 2478; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX coding sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX Sequence 2526 BP; 871 A; 536 C; 541 G; 578 T; 0 U; 0 Other;
```

Alignment Scores:

Pred. No.:	6.32e-25	Length:	2526
Score:	550.00	Matches:	121
Percent Similarity:	66.53%	Conservative:	40
Best Local Similarity:	50.00%	Mismatches:	51
Query Match:	20.36%	Indels:	30
DB:	5	Gaps:	7

US-10-023-523-8 (1-530) x AAS66674 (1-2526)

QY 167 SerThrProGluGluLysLeuAlaLeuCys-----LysLysTyraLalaGluLeu 184

DB 157 AGCAATATGGAAGAA-----GCTGGACTTTGGGGTTAAGAGAGAAAGCAGATATGTTG 210

QY 185 -----GluGluHisArgAsnSerGlnLysGlnMetLysLeuGlnLysGln 201

1

1

QY 109 Pro---ThrProValValTyrGlyGluLysGluProSerLysGlyAspProAsnThrGlu 127
 Db 258 CCAGAGACTCCAGTAGTCAATGGTGAAGAAGAAATCTCCAAGGGGAGCGCGCCGGAC 317
 QY 128 GluileArgGlnSerAspGluValGlyAspArgAsnHisArgArgProGlnGluLysLys 147
 Db 318 GAGATCCGGACCACTGATGAGTCTGAGACCGAGACCACCGAAGGCCACAGGANAAGAAA 377
 QY 148 LysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThr 165
 Db 378 AAGCCAAAGGCTCTGGAAAGGATCAAC-----TGCCCCACTTAATCTACC 425

RESULT 40

ACH31687/c
 ID ACH31687 standard; cDNA; 334 BP.

XX AC ACH31687;

XX 13-OCT-2003 (first entry)

XX Human bone marrow cDNA #978.

XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
 KW genome mapping; biodiversity; genetic disorder.

XX Homo sapiens.

OS US2003073623-A1.

PN 17-APR-2003.

XX 30-JUL-2001; 2001US-00918995.

XX 30-JUL-2001; 2001US-00918995.

XX (DRMA/) DRMANAC R T.

PA (LABA/) LABAT I.

PA (STAC/) STACHE-CRAIN B.

PA (DICK/) DICKSON M C.

PA (JONE/) JONES L W.

XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
 PI WPI; 2003-615964/58.

XX New polynucleotide sequences obtained from various cDNA libraries, useful
 PT as hybridization probes, as oligomers for PCR, for chromosome and gene
 PT mapping, in the recombinant production of protein, or in generating
 PT antisense DNA or RNA.

PS Claim 1; SEQ ID NO 18999; 44pp; English.

XX The invention relates to an isolated polynucleotide comprising any one of
 CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
 CC determined by the technique of SBH (sequencing by hybridisation). Also
 CC included is a purified polypeptide comprising a sequence corresponding to
 CC a reading frame of the novel polynucleotide. The nucleic acid sequences
 CC are useful in diagnostics as expressed sequence tags (EST) for
 CC identifying expressed genes or for physical mapping of the human genome,
 CC in forensics, in assessing biodiversity, or in identifying mutations
 CC responsible for genetic disorders and other traits. The nucleotide
 CC sequences are also useful as hybridisation probes, as oligomers for PCR,
 CC for chromosome and gene mapping, in the recombinant production of
 CC protein, or in generating antisense DNA or RNA. The purified polypeptide
 CC is useful for generating antibodies specific for it. The present sequence
 CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
 CC for this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?docID=20030073623

XX Sequence 334 BP; 55 A; 82 C; 73 G; 124 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.41e-20 Length: 334
 Score: 464.00 Matches: 87
 Percent Similarity: 91.89% Conservative: 15
 Best Local Similarity: 78.38% Mismatches: 9
 Query Match: 17.17% Indels: 0
 DB: 8 Gaps: 0

US-10-023-523-8 (1-530). x ACH31687 (1-334)

QY 210 HisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCys 229
 Db 334 CACTTGCAGAGTGAACATAGCAAGGCTATCTTGCAAGAGCAAGCTAGAAATCTCTTTGC 275
 QY 230 ArgGluLeuGlnArgHisAsnArgSerLeuLysGluGlyValGlnArgAlaArgGlu 249
 Db 274 AGAGACTTCAGCTCACAATAGACGTTAAGAGGAAATATGCAGAGGACGAGAG 215
 QY 250 GluGluGlnLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGln 269
 Db 214 GAAGAAGAACGACGTAAGAAGCAACTGCACATTTCCAGATTACCTTAATGAATTCAA 155
 QY 270 LeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGluLeu 289
 Db 154 GCCAGCTGGAGCAGCATGACATCCACAGCCAACTCCGACAGGAAACATTTGAGCTG 95
 QY 290 AlaGluArgLeuLysLysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAspLys 309
 Db 94 GGGGAGAAGCTAAAGAAGCTCATCGAACAGTACGACTGAGCGGAGGACCATTTGATAAG 35
 QY 310 ValPheLysHisLysAspLeuGlnGlnLeu 320
 Db 34 GTGTTTCAACATAAGGACCTGCAACAGCAGCTC 2

RESULT 41

AAH07907
 ID AAH07907 standard; cDNA; 710 BP.

XX AC AAH07907;

XX 26-JUN-2001 (first entry)

XX Human cDNA clone (5'-primer) SEQ ID NO:4742.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 OS Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-00116126.

XX 29-JUL-1999; 99JP-00248036.

XX 27-AUG-1999; 99JP-00300253.

XX 11-JAN-2000; 2000JP-00118776.

XX 02-MAY-2000; 2000JP-00183767.

XX 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.

XX Claim 1; SEQ ID NO 4742; 2537pp + Sequence Listing; English.

Query Match: 12.95% Indels: 6
DB: 4 Gaps: 2
US-10-023-523-8 (1-530) x AAS35582 (1-405)
Qy 57 GluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGln 76
Db 75 AAAGAACGGGGCTGCCAACTGGAAGACATCTGAGCACATCTGTGTGGACAATAACCG 134
Qy 77 GlyGlyProGlyGluAspGlyValGlnGlyGluProAlaGluProGluAspAlaGluLys 96
Db 135 GGGGGCCCCGGAGGATGGGGACAGGTGAGCCGCTGAACCCGAAGATGCAGAGAAG 194
Qy 97 SerArgThrTyrValAlaArgAsnGlyGluPro-GluProThrProVal---ValTyrGln 115
Db 195 TCCCGGACCTATGTGGCAAGGAATGGGAGCCTTGAACCACTTCCAAGTTCATTGG 254
Qy 115 YGluLysGluProSerIysGlyAsp-ProAsnThrGluGluIle-ArgGlnSer-AspGln 134
Db 255 AGAAGAGGAACCTTCAAGGGGATTCGAACACACAGAGAGATTCGGGACAGATTGACGA 314
Qy 134 uValGlyAspArgAspHis---ArgArgProGlnGluLysLysValAlaLysGlyLeuGln 153
Db 315 AGTTCGGAGCAAGGAACCACTTCGAGGCCACAGAGAGAGAGAAACCCAAAGGTTTGGG 374
Qy 153 YLysGluIleThrLeuLeuMet 160
Db 375 GAGAGATCATCGTTTGTCTGATT 396
RESULT 43
AD845661
ID ADE45661 standard; cDNA; 405 BP.
XX AC ADE45661;
XX DT 29-JAN-2004 (first entry)
XX DE Human cardiovascular system related polynucleotide #457.
XX KW Human; cardiovascular system related polypeptide; cancer;
KW proliferative disorder; foetal abnormality; developmental abnormality;
KW haematopoietic disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW inflammation; allergy; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KW cardiovascular disorder; angioecnic disorder; kidney disorder;
KW gastrointestinal disorder; pregnancy-related disorder;
KW endocrine disorder; gene; ss.
XX OS Homo sapiens.
XX PN US2003059908-A1.
XX PD 27-MAR-2003.
XX PF 07-MAR-2002; 2002US-00091504.
XX PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-020515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225113P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 18-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231988P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233084P.
PR 21-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 25-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 26-SEP-2000; 2000US-0234998P.
PR 27-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.

PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251983P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764869.
(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Ruben SM, Barash SC;
P-PSDB; ADE46276.
WPI; 2003-743766/70.
P-PSDB; ADE46276.

New cardiovascular system related polynucleotides and polypeptides,
useful for preventing, treating, or ameliorating a medical condition,
such as cancer of cardiovascular tissues and cancer metastases.

Claim 1; SEQ ID NO 467; 262pp; English.

The invention relates to human cardiovascular system related polypeptides
and the polynucleotides encoding them. The polypeptides, polynucleotides
and antibodies to the polypeptides are useful for diagnosing a
pathological condition or a susceptibility to a pathological condition,
for preventing, treating, or ameliorating a medical condition, such as
cancer of cardiovascular system tissues, proliferative disorders, foetal
and developmental abnormalities, haematopoietic disorders, diseases of
the immune system, AIDS, autoimmune diseases (e.g., rheumatoid
arthritis), inflammation, allergies, neurological disorders (e.g.,
Alzheimer's disease, Parkinson's disease), cognitive disorders,
schizophrenia, asthma, skin disorders (e.g., psoriasis), sepsis,

CC diabetes, atherosclerosis, cardiovascular disorders, angiogenic
CC disorders, kidney disorders, gastrointestinal disorders, pregnancy-
CC related disorders, endocrine disorders and infections. The nucleic acids
CC are also useful for chromosome identification, radiation hybrid mapping
CC or long-range restriction mapping. The polypeptides and polynucleotides
CC may also be used as food additives or preservatives to increase or
CC decrease storage capabilities, fat content or other nutritional
CC components. This sequence represents a human cardiovascular system
CC related polynucleotide of the invention.

SQ Sequence 405 BP; 123 A; 90 C; 126 G; 63 T; 0 U; 3 Other;

Alignment Scores:

Pred. No.: 1.92e-13 Length: 405
Score: 350.00 Matches: 83
Percent Similarity: 79.09% Conservative: 4
Best Local Similarity: 75.45% Mismatches: 17
Query Match: 12.95% Indels: 6
DB: 9 Gaps: 2

US-10-023-523-8 (1-530) x ADE45661 (1-405)

QY 57 GluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnGln 76
DB 75 AAGAAGCGGGCTGCCAACTGGAAGACATACGTGACACATCTGTGTGGCAATAACCG 134
QY 77 GlyGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLys 96
DB 135 GGGGGCCCCCGCGAGGATGGGGGCACAGGGTGGAGCCGCTGAACCGAGATGCAGAGAAG 194
QY 97 SerArgThrTyrValAlaArgAsnGlyGluPro-GluProThrProVal---ValTyrG 115
DB 195 TCCCGGACCTATGTGGCAAGGAATGGGAGCCTTGAACCACTTCCCACTAGTTCAATTGG 254
QY 115 yGluLysGluProSerLysGlyAsp-ProAsnThrGluGluLeu-ArgGlnSer-AspG 134
DB 255 AGAGAAGGAACCTCCAAAGGGGGATCCAAACACAGAGAGAGATCCCGCAGAGTTGACGA 314
QY 134 uValGlyAspArgAspHis---ArgArgProGlnGlnGlyLysLysAlaLysGlyLeuG 153
DB 315 AGTTCGGAGCAGGAGGACCATTCGAGGCCACAGAGAGAGAGAGAGAGAGAGAGAGAG 374
QY 153 yLysGluLeuThrLeuLeuMet 160
DB 375 GAGGAGATCAGCTTGTCTGATT 396

RESULT 44

ABQ55358
ID ABQ55358 standard; cDNA; 725 BP.

XX AC ABQ55358;

XX DT 22-AUG-2002 (first entry)

XX DE Human ovarian antigen HLTJA50 cDNA, SEQ ID NO:1238.

XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
XX ovarian cancer; breast cancer; tumour; reproductive system disorder;
XX infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
XX PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
XX inflammatory condition; immune disorder; blood disorder;
XX cardiovascular disorder; respiratory disorder; neurological disorder;
XX gastrointestinal disorder; urinary system disorder; drug screening;
XX gene therapy; chromosome mapping; forensic analysis;
XX antibody preparation; cytostatic; immunomodulatory; neuroprotective;
XX antiinflammatory; gynaecological; reproductive; gene; ss.

XX Homo sapiens.

XX WO200200677-A1.

XX 03-JAN-2002.

XX

PF 07-JUN-2001; 2001WO-US018569.
 XX
 PR 07-JUN-2000; 2000US-0209467P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Birse CE, Rosen CA;
 XX
 DR WPI; 2002-147878/19.
 DR P-PSDB; ABP42281.
 XX
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
 PT cancer), immune disorders, cardiovascular disorders and neurological
 PT diseases.
 XX
 PS Claim 1; SEQ ID NO 1238; 2922pp; English.
 XX
 CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis), and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents cDNA encoding a human ovarian antigen of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 725 BP; 263 A; 122 C; 147 G; 176 T; 0 U; 17 Other;
 XX
 Alignment Scores:
 Pred. No.: 3,468-11 Length: 725
 Score: 318.00 Matches: 74
 Percent Similarity: 77.05% Conservative: 20
 Best Local Similarity: 60.66% Mismatches: 26
 Query Match: 11.77% Indels: 3
 DB: 6 Gaps: 0
 US-10-023-523-8 (1-530) x ABQ55358 (1-725)
 QY 347 LeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHis 366
 DB 334 TTAAGAAGGACGACAGATTCGAGGCACAAATACGAAACAAATGAAACACGANGAGTACAA 393
 QY 367 LeuLysGlnGlnLeuAlaLeuThrThrGluLysPheGluGluPheGlnAsnThrLeuSer 386
 DB 394 CTAAGACAGCAGCTTCTCTTATATGATAGTTTGAAGAAATCCAGACTACCATGCA 453
 QY 387 LysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysIle 406
 DB 454 AAACCAATGAACTGGTTACACCTTCAGACAGGAATGGAAGATGACAAAGAAAT 513
 QY 407 LysLysLeuGluLysGluThrThrMetTyrArgSerArgTrpGluSerSerAsnLysAla 426
 XX
 Db 514 AAAAAGCTGAAAAGAAACAAATATTTGGCGNACCAAGGGAACCAT-ATAAAGCA 572
 QY 427 LeuLeuGluMetAlaGluGluLysThrValArgAspLysGluLeuGluGlyLeuGlnVal 446
 Db 573 CTTTCAAAATGGCTGAAAGAGGAAACAGTCGCGNATTAAGAGTNCAGAGNCCTTCAATA 632
 QY 447 LysIleGlnArg-LeuGluLysLeuCysArgAla-LeuGlnThrGluArgAsnAspLeuA 466
 Db 633 AAATGGAAACGGGTAGAGAACTTGGCGANGGCTTNTTCAACCGAAGGATGGGTCA 692
 QY 466 sn 466
 Db 693 AT 694
 RESULT 45
 AEN74510
 ID AEN74510 standard; cDNA; 981 BP.
 XX
 AC AEN74510;
 XX
 DT 03-JUL-2002 (first entry)
 XX
 DE Bovine embryonic germ (EG) cell cDNA EST #561.
 XX
 KW Bovine; Bos taurus; EST; expressed sequence tag; totipotency;
 KW development; gene; ss.
 XX
 OS Bos taurus.
 XX
 PN WO200194550-A2.
 XX
 PD 13-DEC-2001.
 XX
 PP 07-JUN-2001; 2001WO-US018576.
 XX
 PR 07-JUN-2000; 2000US-0209874P.
 PR 06-JUN-2001; 2001US-00876143.
 XX
 PA (INFI-) INFINGEN INC.
 XX
 PI Eilertsen KJ, Pfister-Genskow M, Childs L;
 XX
 DR WPI; 2002-351289/38.
 XX
 PT An expressed sequence tag (EST), the expression of which, or its
 PT complementary sequence, in a cell identifies the cell as a
 PT developmentally competent or incompetent cell.
 XX
 PS Example 16; Page 458-459; 584pp; English.
 XX
 CC The present invention describes an expressed sequence tag (EST), where
 CC the EST is an isolated, enriched, or purified nucleic acid sequence
 CC representing all or part of a gene, the expression of which, or its
 CC complementary sequence, in a cell identifies the cell as a
 CC developmentally competent or incompetent cell. Molecules which induce
 CC developmental competence in a cell line are useful for inducing
 CC totipotency in one or more cells. Molecules which induce developmental
 CC incompetence in a cell line are useful for preventing a full term
 CC pregnancy in an animal and inhibiting totipotency. The molecules are also
 CC useful for treating a disease in an animal by inducing development of one
 CC or more cells of the animal into a specific cell type. The present
 CC sequence represents a bovine EST which is given in the exemplification of
 CC the present invention
 XX
 SQ Sequence 981 BP; 195 A; 212 C; 255 G; 214 T; 0 U; 105 Other;
 XX
 Alignment Scores:
 Pred. No.: 9,816-11 Length: 981
 Score: 313.00 Matches: 74
 Percent Similarity: 62.88% Conservative: 9
 Best Local Similarity: 56.06% Mismatches: 46
 Query Match: 11.58% Indels: 4

Db 4227 CAAGAACGTCATGAGCTGGAGAGTCCAAAGCGGCGCCCTGGAGACCCAGATGAGAGAT 4286
Qy 232 uclnArlHisAsnArgSerLeuLysGluGluValGlnArlHisAsnArgSerLeuLysGluGlu 252
Db 4287 GAAGACGAGCTGAAGAGCTGGAGAGCAG-----CTCAAGACCGAGGAGCGCAA 4340
Qy 252 uLysArgLysGluValThr-----SerHisPheGlnValThrLeuAsnAs 267
Db 4341 ACTCGGCTGGAGTCAACATGACGAGCGCTCAAGGGCCAGTTCGAA-----AGGGA 4391
Qy 267 pLysGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLeuArgGlnGlnAsnMe 287
Db 4392 TCTCAAGCCCGGACGAGCAGATGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4451
Qy 287 tGlu-----LeuAlaGluArgLeuLys 294
Db 4452 CGAGTATGAGACGGAACCTGGAAGACGAGCGAAGCAACAGTGCCTGGCAGCTGACGAAA 4511
Qy 294 sLysLeuIleGlu-----GlnTyrGluLeu----- 302
Db 4512 GAAGACGCTGAAGAGGAGACCTGAAGACCTGGAGCTTCAGGCCGACTCTGCCATCAAGGG 4571
Qy 303 -ArgGluGluHisIleAspLysValPheLys-----HisLysAspLeuGln 318
Db 4572 GAG 4631
Qy 318 nGlnLeuValAspAlaLys----- 324
Db 4632 AGAGCTGAAGATGCCCGTGCCTCCAGAGATGAGATCTTCCACACGCCAAAGAGATGA 4691
Qy 325 -----LeuGlnGlnAlaGlnGlnMetLeuLysGluAl 335
Db 4692 GAAGAACCCAGAGCTTGGAGGAGGAGCTATGAGCTACAGAGGACCTCGCGCGCG 4751
Qy 335 aGluGluArgHisGlnArg-----GluLysAspPheLeuLeuLysGluAlaValG 352
Db 4752 TGAGAGGCTCGCAACCAACAGCGGAGCTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4811
Qy 352 uSerGlnArgMetCysGluLeuMetLysGlnGlnThrHisLeuLysGlnGlnLeuAl 372
Db 4812 TAGCTGTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4871
Qy 372 aLeuTyrThrGluLysPheGluGlnPheGlnAsnThrLeuSerLysSerSerGluValPh 392
Db 4872 CCAGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4926
Qy 392 eThrThrPheLysGlnGlnMetGluLysMetThrLysLysLysLysLysLysLysLysG 412
Db 4927 -----CGGTCGCGCAAGCCACACAGCAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4970
Qy 412 uThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeuGluMet-----Al 431
Db 4971 GTTGGCCACAGAGCGAGCAGCGCCAGAGATGAGTGCCTGGCAGGAGCTCGAGCG 5030
Qy 431 aGluGluLysThrValArgAspLys-----GluLeuGluGly----- 443
Db 5031 GCAGAACAGGAGCTCCGAGCAAGCTCCACAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5090
Qy 444 -----LeuGlnValLysIleGlnArgLeuGluLysLeu----- 454
Db 5091 CAAGTCCACCATCGGCGCTGGAGCGCCAGATTCCACAGCTGGAGGAGGAGGAGGAGGAGGAGGAG 5150
Qy 455 ----CysArgAlaLeuGlnThrGluArgAsnAspLeu-----AsnLysArgValG 470
Db 5151 GGAGGCCAGAGAGAACAGCGCGCCACCAAGTCGCTGAAGCAGCAAGAAAGCAAGAGCTGAA 5210
Qy 470 nAspLeuSerAlaGlyGlyGlnGlySerLeuThrAsp-----SerGln 484
Db 5211 GGAATCTT-GCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5269
Qy 484 yProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrG 504
Db 5270 CAGAGAACGCAATGCCAGGCTCAAGCAGCTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5329

Qy 504 uAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGln----- 519
Db 5330 ACTCCAGCGCATCAACCGCAACCGAGAGCTCGAGCGGAGCTGATGAGGCCAGCG 5389
Qy 520 -----ThrGlyPro-----GlnGluProThrSerAlaArg 529
Db 5390 AGAGCAACGAGCGCATGGCGCGGAGGTGAACGCACTCAAGAGCAAGCTCAGCGGAGG 5447
RESULT 47
ACA62126
ID ACA62126 standard; cDNA; 5937 BP.
XX AC
XX ACA62126;
XX
XX 20-AUG-2003 (first entry)
XX Human smooth muscle myosin heavy chain (hSMMyHC) variant cDNA #2.
XX
XX Human; smooth muscle myosin heavy chain; hSMMyHC; inotropic;
KW antiasthmatic; hypertensive; hypotensive; utopathic; antiinflammatory;
KW gynaecological; cololytic; cardiac; vascular disorder;
KW pulmonary disorder; reproductive disorder; immunological disorder;
KW asthma; hypotension; hypertension; urinary incontinence;
KW irritable bowel syndrome; menstrual labor; premature labor; glaucoma;
KW bronchioconstriction; cardiac malfunction; gene; ss.
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 1. 5937
FT CDS
FT FT /*tag= a
FT /product= "hSMMyHC variant"
FT /note= "human smooth muscle myosin heavy chain"
FT /partial
FT /note= "No stop codon given"
XX
XX US2003032018-A1.
XX
XX 13-FEB-2003.
XX
XX 10-AUG-2001; 2001US-00927597.
XX
XX 10-AUG-2001; 2001US-00927597.
XX
XX (CYTO-) CYTOKINETICS INC.
XX
XX Malik F, Beraud C, Freedman R, Craven A, Sakowicz R, Hartman J;
XX
XX WPI; 2003-479585/45.
XX P-PSDB; ABU10399.
XX
XX New human smooth muscle myosin heavy chain (hSMMyHC) polypeptide and
XX nucleic acid, useful for diagnosing, preventing or treating e.g. asthma,
XX hypertension, incontinence, menstrual cramps, premature labor or cardiac
XX malfunction.
XX
XX Claim 4; Fig 3A-D; 80pp; English.
XX
XX The invention describes an isolated human smooth muscle myosin heavy
XX chain (hSMMyHC) polypeptide. The hSMMyHC protein, nucleic acid, or its
XX modulator, is useful for diagnosing, preventing or treating vascular,
XX pulmonary, reproductive or immunological disorders. In particular, the
XX hSMMyHC polypeptide, nucleic acid, or its modulator is useful for
XX diagnosing, preventing or treating e.g. asthma, hypotension,
XX hypertension, urinary incontinence, irritable bowel syndrome, menstrual
XX cramps, premature labor, glaucoma, bronchioconstriction, cardiac
XX malfunction, or other medical conditions related to hSMMyHC function. The
XX hSMMyHC polypeptide or nucleic acid is also useful for screening
XX therapeutic agents or hSMMyHC modulators, which may be used for treating
XX the above-mentioned diseases or disorders. This sequence encodes a human
XX smooth muscle myosin heavy chain (hSMMyHC) variant missing exon 42

RESULT 48
ACF12909

Db 5125 GAAGAAAGCAAGAGCTTGGAGAGCAGACCTCATGAGCTACAGAGGACCTCGCGCGGC 5184
 Qy 335 aGluGluArgHisGlnArg-----GluLysAspPheLeuLeuLysGluAlaValG1 352
 Db 5185 TGAGAGGGCTCGCAACAAAGAGGAGCTCGAGAGGAGGAACTGGCAGAGGAGCTGGCCAG 5244
 Qy 352 uSerGlnArgMetCysGluLeuMetLysGlnGlnThrHisLeuLysGlnGlnLeuAl 372
 Db 5245 TAGCCCTGTGGGAGAAAGCGCTCCAGAGCAGAGAGCCGCCCTGGAGGCCGATCGC 5304
 Qy 372 aLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluValph 392
 Db 5305 CCAGCTGGAGGAGGAGCTGGAGGAGGAGCAGGCAACATGGAGGCCATGAGCGAC----- 5359
 Qy 392 eThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLeuLysG1 412
 Db 5360 -----CGGGTCGCAAAAGCCACAGCAGGAGCGCGGAGCTCAGCAACGA 5403
 Qy 412 uThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeuGluMet---Al 431
 Db 5404 GTGGCCACAGAGCGCAGCAGCGCCAGAGATGAGTGCCTGGAGGAGCGCTCGACG 5463
 Qy 431 aGluGluLysThrValArgAspLys-----GluLeuGluGly----- 443
 Db 5464 GCAGAACAGAGGAGCTCGGAGCAAGCTCCAGAGATGGAGGGGCGCGTCAAGTCCAGTT 5523
 Qy 444 -----LeuGlnValLysLysGlnArgLeuGluLysLeu----- 454
 Db 5524 CAAGTCCACATCGCGGCGCTGGAGCCCAAGATTGCACAGCTGGAGGAGCGAGTGCAGCA 5583
 Qy 455 ----CysArgAlaLeuGlnThrGluArgAsnAspLeu-----AsnLysArgValG1 470
 Db 5584 GGAGGCCAGAGAGAAACAGCGCGCCACCAAGTCGTGAAGCAGAGAGAAAGAGCTCAA 5643
 Qy 470 nAspLeuSerAlaGlyGlyGlnGlySerLeuThrAsp-----SerG1 484
 Db 5644 GAAATCTTT-GCTGCAGGTGGAGGAGCAGCGCAAGATGCGCAGCAGTACAGAGGAGCAG 5702
 Qy 484 yProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrG1 504
 Db 5703 CAGAGAGGCGAATCCAGGCTCAGCAGCTCAGAGCAGCTGAGGAGGAGGAGG 5762
 Qy 504 uAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGln----- 519
 Db 5763 AGTCCGAGGCGCATCAACGCCAACCGCAGGAGCTGCAGCGGAGCTGGATGAGGCCACGG 5822
 Qy 520 -----ThrGlyPro-----GlnGluProThrSerAlaArg 529
 Db 5823 AGAGCAACAGAGCCATGGGCCCGGAGGTGACCGACTCAAGAGCAGCTCAGGCGAGG 5880
 RESULT 49
 ACH03809/c
 ID ACH03809 standard; cDNA; 11065 BP.
 XX AC ACH03809;
 XX DT 26-SEP-2003 (first entry)
 XX DE Human cDNA differentially expressed in lung cancer #14.
 XX KW Gene therapy; emphysema; ss; gene; chronic obstructive pulmonary disease;
 XX KW respiratory disorder; lung cancer; asthma; human.
 XX OS Homo sapiens.
 XX PN US2003065157-A1.
 XX PD 03-APR-2003.
 XX PF 04-APR-2002; 2002US-00116802.
 XX PR 04-APR-2001; 2001US-0281593P.

XX PA (LASE/) LASEK A W.
 XX PI Lasek AW;
 XX DR WPI; 2003-540803/51.
 XX PT New combination comprising cDNAs that are differentially expressed in
 PT respiratory disorders, useful for diagnosing or treating respiratory
 PT disorders e.g., lung cancer, chronic obstructive pulmonary disease,
 PT emphysema or asthma.
 XX PS Claim 1; Page; 39pp; English.
 XX CC The invention relates to a combination comprising cDNAs or their
 CC complements that are differentially expressed in respiratory disorder.
 CC The combination is useful for preparing a composition for diagnosing or
 CC treating respiratory disorders e.g. lung cancer, chronic obstructive
 CC pulmonary disease, emphysema or asthma. The present sequence represents
 CC human cDNA differentially expressed during lung cancer
 XX SQ Sequence 11065 BP; 2564 A; 3049 C; 3072 G; 2377 T; 0 U; 3 Other;
 Alignment Scores:
 Pred. No.: 1.07e-07 Length: 11065
 Score: 282.50 Matches: 153
 Percent Similarity: 38.48% Conservative: 101
 Best Local Similarity: 23.18% Mismatches: 229
 Query Match: 10.46% Indels: 178
 DB: 8 Gaps: 27
 US-10-023-523-8 (1-530) x ACH03809 (1-11065)
 Qy 3 SerProGlyGlnProGluAla-----GlyProGluGlyAlaGlnGluArgPro 18
 Db 9685 GCTCCAGGACCCAGGAGCTGCTTCAAGAGAAACCGCAGAGCTCAA---CGTGTG 9629
 Qy 19 SerGlnAlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLys 38
 Db 9628 TAGCAAGCTGCGCAGCT-----GGAGGAGGAGCGGAGCAAGCTCAAGA 9584
 Qy 39 ProGluGly---AlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGlu 57
 Db 9583 CCAGCTGAGCAGGAGATGGAGCCAGCAGACCTGGAGCGCCACATCTCCACCTCAA 9524
 Qy 58 GluLeuSerArgGlnLeuGluAspLysLeuSerThrTyrCysValAspAsnGlnGly 77
 Db 9523 CATCCAGCTCTCCGACTCGAAGAGAAAGAGCTGCAGGAGCTTTGCCAGCACCGTGAAGCTCT 9464
 Qy 78 GlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluPro----- 91
 Db 9463 GGAAGAGGGGAGAGAGAGGTTCCAGAGAGAGATCGAGAACCTCACCCAGCAGTACAGGA 9404
 Qy 92 -----GluAspAlaGluLysSerArgThrTyrValAla 102
 Db 9403 GAAGCGCGCGGTATGATAAACTGGAAGAACACCAAGAACAGCTTCAGCAGAGCTGA 9344
 Qy 103 ArgAsn-----GlyGluProGluProThrProValValTyrGlyGluLysGlu 118
 Db 9343 CGACCTGGTTGTTGATTGGACAACCGCGCAACTCGTGTCAACCTCGAAGAGAGCA 9284
 Qy 119 Pro-----SerLysGlyAspProAsnThrGluGluLeuArgGlnSer 132
 Db 9283 GAGGAATTTGATCAGTTGTTAGCCGAGGAGAAACATCTCTTCCAAATACGC----- 9230
 Qy 133 AspGluValGly-AspArgAspHisArgProGlnGluLysLysLysAlaLysGlyLe 152
 Db 9229 -----GGATGAGAGGACAGAGCTGAGGAGAGAGCGAGGAGGAG- 9190
 Qy 152 uGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLys 172
 Db 9189 -----GAAACCAA 9182

Percent Similarity: 38.33% Conservative: 99
 Best Local Similarity: 23.33% Mismatches: 230
 Query Match: 10.42% Indels: 178
 DB: 9 Gaps: 26

US-10-023-523-8 (1-530) x ADB85060 (1-3320)

QY 3 SerProGlyGlnProGluAla-----GlyProGluGlyAlaGlnGluArgPro 18
 DB 676 GCTCCAGACACCCAGGAGCTGCTTCAAGAGAAACCCGCGCAGAACTCAA---CGTGTC 732
 QY 19 SerGlnAlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLys 38
 DB 733 TACGAAGCTGGCCAGCT-----GGAGAGAGCGGAAACAGCTGCAAGA 777
 QY 39 ProGluGly---AlaGlnAlaArgThraGlnSerGlyAlaLeuArgAspValSerGlu 57
 DB 778 CCAGCTGACGAGAGATGGAGGCCAAGCAGAACCTCGAGCGCCACATCTCCACTCTCAA 837
 QY 58 GluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGly 77
 DB 838 CATCCAGCTCTCCACTCGAAGAGAGAGCTGCGAGACTTTGCCAGCACCGCTGAAAGCTCT 857
 QY 78 GlyProGluAspAlaGlnGlyGluProAlaGluPro----- 91
 DB 898 GGAAGAGCGGAGAGAGAGGTTCAGAGAGAGATCGAGAACCTCACCCAGCAGTACGAGGA 957
 QY 92 -----GluAspAlaGluLysSerArgThrTyrValAla 102
 DB 958 GAAGGCGCGCGGTATGATATAACTGGAAGAGACCAAGAACAGAGCTTCAGCAGAGCTGGA 1017
 QY 103 ArgAsn-----GlyGluProGluProThrProValValTyrGlyGluLysGlu 118
 DB 1018 CGACCTGTTGTTGATTGGAACACACGAGCGCACTCGTGTCACACTGGAAAGAGCA 1077
 QY 119 Pro-----SerLysGlyAspProAsnThrGluGluIleArgGlnSer 132
 DB 1078 GAGGAATTTGATCAGTTGTTAGCCGAGGAGAGAAACATCTCTTCCAAATACGC----- 1131
 QY 133 AspGluValGly-AspArgAspHisArgProGlnGluLysLysAlaLysGlyLe 152
 DB 1132 -----GGATGAGAGGACAGACTGAGGAGAGAGCCAGGAGAG- 1171
 QY 152 uGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGlu 172
 DB 1172 -----GAAACCAA 1179
 QY 172 sLeuAlaAlaLeuCysLysLysTyrAlaGluLeuLeuGluHisArgAsnSerGlnly 192
 DB 1180 GGCCCTGTCCTCGCTCGGCGCCCTTGAAGAGCCCTTGAAGCCAAAGAGAGAACTCGAGCG 1239
 QY 192 sGlnMethLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuAr 212
 DB 1240 GACCAACAAATGCTCAAAGCCGAATGAAGACCTGGTCACTCCAGGATGACGTGGG 1299
 QY 212 gGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLe 232
 DB 1300 CAAGAACGTCATGAGCTGGAGAAATCCAAAGCGGCCCTGGAGACCCAGATGGAGAGAT 1359
 QY 232 uGlnArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGluGluGlu 252
 DB 1360 GAAGAGCGAGCTGGAAGACTGGAGGAG-----CTGCAGACCCAGGAGGCCCAA 1413
 QY 252 uLysArgLysGluValThr-----SerHisPheGlnValThrLeuAsnAs 267
 DB 1414 ACTGGCGCTGGAAGTCAACATGTCAGGCGCTCAAGGCGCCAGTTCGAA-----AGGA 1454
 QY 267 pIleGlnLeuMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMe 287
 DB 1465 TCTCCAGCCCGGACGAGCAGATGATGAGAGAGAGGAGGAGCACTGCGAGAGAGCTTCA 1524
 QY 287 tGlu-----LeuAlaGluArgLeuLy 294

DB 1525 CGAGTATGAGACGGAACCTGGAAGACGAGCGAAAGCAACGCTGCCCTGGCAGCTGCAGCAA 1584
 QY 294 sLysLeuIleGlu-----GlnTyrGlnLeu----- 302
 DB 1585 GAAGAAGCTTGAAGAGGGGACCTGAAAGACCTGGAGCTTCAGGCCGACTCTGCCATCAAGGG 1644
 QY 303 -ArgGluGluHisIleAspLysValPheLys-----HisLysAspLeuGlnG 318
 DB 1645 GAGGAGGAAGGCAATCAAGCAGCTACGCAAACTGCAGGCTCAGATGAGAGGACTTTCAGAG 1704
 QY 318 nGlnLeuValAspAlaLys----- 324
 DB 1705 AGAGCTGGAAGATGCCCGTGCCTCCAGAGATCAGATCTTTGCCACAGCCAAAGAGAATGA 1764
 QY 325 -----LeuGlnGlnAlaGlnGluMetLeuLysGluAl 335
 DB 1765 GAAGAAAGCCAGAGCTTGGAGCAGACCTCATGCAGCTACAAGAGGACCTTCGCCGCCGC 1824
 QY 335 aGluGluArgHisGlnArg-----GluLysAspPheLeuLeuLysGluAlaValG 352
 DB 1825 TGAGAGGCTCGCAACCAAGCGGACCTCGAGAGGAGGAACTGGCAGAGGAGCTGGCCAG 1884
 QY 352 uSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLeuLysGlnGlnLeuAl 372
 DB 1885 TAGCTGTCCGGAAGAACGCACTCCAGGACGAGAGCGCGCTGGAGGCCCGGATCGC 1944
 QY 372 aLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluValPh 392
 DB 1945 CCAGCTGAGGAGAGCTGGAGGAGGAGGAGGCAACATGAGGAGCCATGAGCGAC----- 1999
 QY 392 eThrThrPheLysGlnGluMetGluLysMetThrLysLysIleLysLysLeuGluLysG 412
 DB 2000 -----CGGTCCCAAAAGCCACACAGCGGCCGCGAGCAGCTCAGCAACGA 2043
 QY 412 uThrThrMetTyrArgSer-----ArgTyrGluSerSerAsnLysAlaLeuLeuG 429
 DB 2044 GCTGCCACAGAGCGCAGCAGCGCCAGAGAGATGAGATTCCTCCGCGCAGCAGCTTGAGCG 2103
 QY 429 uMetAlaGluGluLysThrValArgAspLysGluLeuGluGly----- 443
 DB 2104 GCAGAACAGAGAGCTCCGAGCAGAGCTCCACGAGATGGAGGGGCGCTCAAGTCCAAGTT 2163
 QY 444 -----LeuGlnValLysIleGlnArgLeuGluLysLeu----- 454
 DB 2164 CAATCCACCATCGCGCGCTGGAGGCCAAGATTCACAGCTAGAGGAGCAGCTCAGCA 2223
 QY 455 -----CysArgAlaLeuGlnThrGluArgAsnAspLeu-----AsnLysArgValG 470
 DB 2224 GGAGGCCAGAGAGAAAACAGCGCGCCACCAAGTCCGCTGAAGCAAGAAAGAGCTGAA 2283
 QY 470 nAspLeuSerAlaGlyGlyGlnGlySerLeuThrAsp-----SerG 484
 DB 2284 GGAATCTT-GCTGCAGTGGAGAGCAGCGCGAAGATGGCCGAGCAGTACAGGAGCAGG 2342
 QY 484 YProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrG 504
 DB 2343 CAGAGAAAGCAATGCCAGGCTCAAGCAGCTCAAGAGGAGCTGGAGGAGGAGGAGGAGG 2402
 QY 504 uAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGln----- 519
 DB 2403 AGTCCAGCGCATCAACGCCAACCCGAGGAAGCTGCGAGGGAGCTGGATAGGCCACCG 2462
 QY 520 -----ThrGlyPro-----GlnGluProThrSerAlaArg 529
 DB 2463 AGAGCAACGAGGCCATGGCGCGGAGGTGCAACGCACTCAAGAGCAGCTCAGGCCGAGG 2520

Search completed: June 8, 2004, 07:01:41
 Job time : 796.959 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus p2n model

Run on: June 8, 2004, 06:37:54 ; Search time 6051.65 Seconds
(without alignments)
3795.955 Million cell updates/sec

Title: US-10-023-523-8

Perfect score: 2702

Sequence: 1 KSSPGQPEAGPEGAQRPSPQ.....APSTEASGQTGPQPTSARA 530

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 70.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 50 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cpn2.1/USPTO.spool/US10023523/runat.08062004.063748.29937/app.query.fasta.1.1422
-DB=GenEmb1 -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=50
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=50 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10023523 @CN 1.1.7704 @runat.08062004.063748.29937 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOF=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmb1 :

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sv.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pin.*
35: em.htg.rod.*
36: em.htg.man.*
37: em.htg.vrt.*
38: em.sv.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2702	100.0	4697	6	AR199536 Sequence
2	2702	100.0	4697	6	AR374688 Sequence
3	2702	100.0	4697	6	AR409323 Sequence
4	2702	100.0	4697	6	AX239574 Sequence
5	2702	100.0	4697	6	BD056450 Novel low
6	2693	99.7	1638	6	AR409338 Sequence
7	2693	99.7	1638	6	AX239603 Sequence
8	2693	99.7	1993	9	AF516206 Homo sapi
9	2689	99.5	4880	9	AL832636 Homo sapi
10	2688	99.5	4851	9	AL832637 Homo sapi
11	2649	98.0	5023	9	HSMB03646
12	2465.5	91.2	4722	6	AR199533 Sequence
13	2465.5	91.2	4722	6	AR374685 Sequence
14	2465.5	91.2	4722	6	AR409320 Sequence
15	2465.5	91.2	4722	6	AX239571 Sequence
16	2465.5	91.2	4722	6	BD056447 Novel low
17	2382	88.2	4702	10	BC060327 Mus muscu
18	1727	63.9	1793	6	AX578043 Sequence
19	1688	62.5	1793	9	HL5344 Human high
20	1452	53.7	181547	2	AC108815 Mus muscu
21	1285	47.6	2045	6	AX878120 Sequence
22	1285	47.6	2045	6	BD157034 Primer fo
23	1285	47.6	2045	9	AK002071 Homo sapi
24	1286.5	46.9	4625	9	BSX647633 Homo sapi
25	1286.5	46.9	4640	9	AL831988 Homo sapi
26	1285.5	46.8	2869	9	AB085905 Homo sapi
27	1265.5	46.8	4631	9	HSMB05276
28	1260	46.6	242048	2	AC110537 Mus muscu
29	1258.5	46.6	4647	9	HSMB07623
30	1240	45.9	4704	5	D89999 Gallus gall
31	1240	45.9	4723	6	E26533 Protein wit
32	1229.5	45.5	2435	10	AF422244 Mus muscu
33	1229.5	45.5	4394	10	AF422245 Mus muscu
34	1054	39.0	1226	10	BC059099 Mus muscu
35	1047.5	38.8	1191	6	AR072731 Sequence
36	1047.5	38.8	1191	6	BD066992 Modulator
37	1008	37.3	4154	9	AL832322 Homo sapi
38	925.5	34.3	823	6	AX053282 Sequence
39	820.5	30.4	2194	6	AX834462 Sequence
40	820.5	30.4	2194	9	AK096939 Homo sapi
41	771	28.5	141790	10	AL671759 Mouse DNA
42	721.5	26.7	22255	6	AR409342 Sequence
43	721.5	26.7	22255	6	AX239608 Sequence
44	713	26.4	529	6	BD060509 Secreted
45	707	26.2	3554	9	BC045655 Homo sapi
46	682	25.2	2181	3	AK117024 Ciona int
47	671	24.8	115756	9	BC029686 Homo sapi
48	666	24.6	3495	9	BC029686 Homo sapi
49	597.5	22.1	1841	3	AY047502 Drosophil
50	534	19.8	245095	2	AC111774 Rattus no

ALIGNMENTS

[illegible]

```

AUTHORS      Lees,R.S., Law,S.W. and Arjona,A.A.
TITLE        Low density lipoprotein binding proteins and their use in
              diagnosing and treating atherosclerosis
JOURNAL      Patent: US 6355451-A 17 12-MAR-2002;
FEATURES     Location/Qualifiers
              1..4697
              /organism="unknown"
              /mol_type="unassigned DNA"

ORIGIN

Alignment Scores:
Pred. No.:      1.79e-124      Length:      4697
Score:          2702.00      Matches:     530
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches:  0
Query Match:    100.00%      Indels:      0
                6              Gaps:        0
                DB:

US-10-023-523-8 (1-530) x AR199536 (1-4697)

```

QY	1	LysSerSerProGlyGluProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln	20
Db	3	AAAAGCAGCCACGAGCAACCCGAGACGAGACCCGAGGAGCGCCACGACGAG	62
QY	21	AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu	40
Db	63	GCGGCTCTTGCAAGTAAAGACGAAAGGTCCTCCGACAGCCAGGCTCTCTGGAAAGCCGAG	122
QY	41	GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer	60
Db	123	GCGGCTCAAGCCAGAACCGCTCAGTCTGGGCGCTCTGTCATGTCTCTGAGGAGCTGAGC	182
QY	61	ArgGlnLeuGluAspIleLeuSerThrTyrcysValaspAsnAsnGlnGlyGlyProGly	80
Db	183	CGCCAACTGGAGACATATCTAGACACATCTGTGTGACAAATAACACAGGGGGGGCCCCGGC	242
QY	81	GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr	100
Db	243	GAGGATGGGCACAGGGTGACCGGCTCAACCCGAGATGCACAGAGATCCCGACCTAT	302
QY	101	ValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGluLysGluProSer	120
Db	303	GTGGCAAGGAATGGGGAGCCCTGAACCAACTCCAGTAGTCTATGGAGAGAGAACCCCTCC	362
QY	121	LysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHis	140
Db	363	ANGGGGGATCCAAACACAGAGAGATCCGGCAGAGTGACCGAGTCCGAGACCCGAGACCAT	422
QY	141	ArgArgProGlnGluLysLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMet	160
Db	423	CGAAGCCACACGAGAGAAAAAAGCCAAAGGGTTGGGGAAAGGAGATCACGTGTCTGATG	482
QY	161	GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysIleTyr	180
Db	483	CAGCATTTGAATACTCTGAGTACCCACAGAGAGAGCTGGCTGCTCTGTGCAGAGAAAGTAT	542
QY	181	AlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLys	200

DEFINITION	Sequence 17 from patent US 6605588.
ACCESSION	AR374688
VERSION	AR374688.1
GI	40077503
KEYWORDS	Unknown.
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 4697)
TITLE	Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.
JOURNAL	Low density lipoprotein binding proteins and their use in
FEATURES	diagnosing and treating atherosclerosis
source	Patent: US 6605588-A 17 12-AUG-2003;
	Location/Qualifiers
	1..4697
	/organism="unknown"
	/mol_type="genomic DNA"
ORIGIN	
Alignment Scores:	
Pred. No.:	1.79e-124 Length: 4697
Score:	2702.00 Matches: 530
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	100.00% Indels: 0
DB:	Gaps: 0
US-10-023-523-8 (1-530) x AR374688 (1-4697)	
Qy	1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 20
Db	3 AAAGCGCCAGCCAGGACACCGGAAGCAGGACCCGAGGAGCCCGAGGAGCGGCCACGCCAG 62
Qy	21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40
Db	63 GCGGCTCTCGAGTAGACGACGAGAGTCCCGGACGAGCCAGGCTCTCGGAAGCCGCGAG 122
Qy	41 GlyAlaGlnAlaArThrAlaGlnSerGlyAlaLeuArgaspValSerGluGlnLeuSer 60
Db	123 GGGGCTCAAGCCAGAACCGCTCAGTCTGGGCCCCCTCGTGATGCTCTCGAGGAGGTGAGC 182
Qy	61 ArgGlnLeuGluAspIleuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGly 80
Db	183 CGCCAACTGGAGACATCTAGACACATCTGTGTGGACAAATACACAGGGGGCCCCGGC 242
Qy	81 GluAspGlyValaGlnGlyGluProAlaGluProGluAlaGluSerArgThrTyr 100
Db	243 GAGGATGGGCGCACAGGGTGAGCGGCTGAACCCGAGATGCAGAGAAGTCCCGACCTAT 302
Qy	101 ValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGluLysGluProSer 120
Db	303 GTGGCAAGGAATGGGAGCCCTGAACCAACTCCAGTAGTCTATGAGAGGAAGAACCTCC 362
Qy	121 LysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHis 140
Db	363 AAGGGGGATCCAAACACAGAAGAGATCCGGCAGAGTGACGAGGTCCGAGACCGAGACCAT 422
Qy	141 ArgArgProGlnGlnLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMet 160
Db	423 CGAAGGCCACAGGAGAAGAAAAGCCAAAGGTTTGGGGAGGAGATCACGTTGCTGATG 482
Qy	161 GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysTyr 180
Db	483 CAGACATTGAATACTCTGATACCCACAGGAGAGCTGGCTGCTCTGTGCAAGAGTAT 542
Qy	181 AlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLys 200
Db	543 GCTGAATCTGTGGAGGAGCACCGGAATTCACAGAAGCAGAGATGAAGTCTCTACAGAAAAG 602
Qy	201 GlnSerGlnLeuValGlnGlnLysAspHisLeuArgGlyGluHisSerLysAlaValLeu 220
Db	603 CAGAGCCAGTGGTGCAAGAGAAGGACCACTTGCCTGAGCACACAGAGGCCGCTCTG 662
Qy	221 AlaArgSerLysLeuGluSerLeuLysCysArgGluLeuGlnArgHisAsnArgSerLeuLys 240

REFERENCE 1 (bases 1 to 4697)
 AUTHORS Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.
 TITLE Low density lipoprotein binding proteins and their use in
 diagnosing and treating atherosclerosis
 JOURNAL Patent: US 6632923-A 17 14-OCT-2003;
 FEATURES Location/Qualifiers
 source 1..4697

ORIGIN
 /organism="unknown"
 /mol_type="genomic DNA"

Alignment Scores:
 Pred. No.: 1,79e-124 Length: 4697
 Score: 2702.00 Matches: 530
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-023-523-8 (1-530) x AR409323 (1-4697)

Qy 1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 20
 Db 3 AAAGAGCCCGAGGACCAACCGGAGCAGAGCCGAGGAGCCCGAGGAGCCCGAGGAG 62
 Qy 21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40
 Db 63 GGGGCTCTGTCAGTAGAGGAGGAGGAGGTCCTCGGAGCCGAGGAG 122
 Qy 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluLeuSer 60
 Db 123 GGGGCTCAAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 182
 Qy 61 ArgGlnLeuGluAspIleLeuSerThrTyrcysValAspAsnGlnGlyGlyProGly 80
 Db 183 CCGCAACTGGAGAGACATCTGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 242
 Qy 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
 Db 243 GAGGATGGGGACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 302
 Qy 101 ValAlaArgcAsnGlyGluProGluProThrProValValTyrGlyGluLysGluProSer 120
 Db 303 GTGGCAAGGAATGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 362
 Qy 121 LysGlyAspProAsnThrGluGluAlaArgGlnSerAspGluValGlyAspArgAspHis 140
 Db 363 AAGGGGATCCAAACACAGAGAGAGATCCGGCAGAGTACGAGGAGGAGGAGGAGGAGGAG 422
 Qy 141 ArgArgProGlnGluLysGlyLysAlaLysGlyLeuGlyLysGluLysGluLysLeuMet 160
 Db 423 CGAAGGCCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 482
 Qy 161 GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuLysLysLysTyr 180
 Db 483 CAGACATTGAATCTCTGAGTACCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 542
 Qy 181 AlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLys 200
 Db 543 GCTGAACCTGTGGAGGAGCAGCGGAATTCACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 602
 Qy 201 GlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeu 220
 Db 603 CAGAGCCAGCTGTGGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 662
 Qy 221 AlaArgSerLysLeuGluSerLeuLysCysArgGluLeuGlnArgHisAsnArgSerLeuLys 240
 Db 663 GCCCGCAGCAAGCTTGAGAGGCTATCGCGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 722
 Qy 241 GluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHis 260
 Db 723 GAAGAAGGTGTGCAGCGGGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 782

Qy 261 PheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisGlnGluArgHisSer 280
 Db 783 TTCAGGTGACACTGAATGACATTCAGCTGCAGATGGAACAGACACATGAGCGCACTCC 842
 Qy 281 LysLeuArgGlnGlnAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTyr 300
 Db 843 AAGCTGGCCCAAGAGAACATGAGGCTGGCTGAGAGGCTCAAGAAGCTGATTGAGCAGTAT 902
 Qy 301 GluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeu 320
 Db 903 GAGCTGGCGAGGAGCATATCGACAAAGTCTTCAAACACAGAGGACCTACAAAGCAGCTG 962
 Qy 321 ValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluGluArgHisGln 340
 Db 963 GTGGATGCCAAGCTCCAGCAGGCCCGCAGGAGATGCTAAAGAGGCGCAGAGGGGACCCAG 1022
 Qy 341 ArgGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMet 360
 Db 1023 CCGGAGAGGAGTCTTCTCTGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1082
 Qy 361 LysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGlu 380
 Db 1083 AAGCAGCAAGAGACCCACCTGAGCAACAGCTTGCCCTATACACAGAGAGGTTGAGGAG 1142
 Qy 381 PheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGlu 400
 Db 1143 TTCAGAACACACTTTCCAAAGAGCAGGAGGTATTCCACCACATTCAAGCAGGAGATGAA 1202
 Qy 401 LysMetThrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 420
 Db 1203 AGATGACTAAGAGATCAAGAGCTGGAGAAAGAACACCATGTACGGTCCCGGTGG 1262
 Qy 421 GluSerSerAsnLysAlaLeuLeuMetAlaGluGluLysThrValArgAspLysGlu 440
 Db 1263 GAGAGCAGCAACAGGCCCTCTGAGATGGCTGAGGAGAAACAGTCCGGGATTAAGAA 1322
 Qy 441 LeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuLysCysArgAlaLeuGlnThr 460
 Db 1323 CTGGAGGGCTCGAGGTAAATCCAAACGCTGGAGAAAGCTGTGCCGGCAGCTGCAGACA 1382
 Qy 461 GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeu 480
 Db 1383 GAGCGCAATGACTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1442
 Qy 481 ThrAspSerGlyProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerPro 500
 Db 1443 ACTGACAGTGGCCCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1502
 Qy 501 ArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThr 520
 Db 1503 AGGGTCACAGAGCGGCTTGTACCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1562
 Qy 521 GlyProGlnGluProThrSerAlaArgAla 530
 Db 1563 GGGGCTCAAGAGCCCACTCCCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1592

RESULT 4

AX239574 4697 bp DNA linear PAT 26-SEP-2001
 Sequence 17 from Patent WO0164874.
 DEFINITION AX239574
 ACCESSION AX239574
 VERSION AX239574.1 GI:15797259
 KEYWORDS Homo sapiens (human)
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.
 Low density lipoprotein binding proteins and their use in diagnosing
 and treating atherosclerosis
 Patent: WO 0164874-A 17 07-SEP-2001;
 Boston Heart Foundation, Inc. (US)

FEATURES
source Location/Qualifiers
1. 4697
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
<3_1595
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC88297.1"
/db_xref="GI:115797260"
/db_xref="RENTREMBL:CAC88297"
/translation="KSPSGPEAGPEGAQRPQAAPAVEAGPGSSQAQPRKPEGAQA
RTAQSGALRVSELSRQLSDILSTVCDNNOGPGEDGAGQGPAPEDAKSGRTVVA
RNEPPTPVVYGKEPSKDPNTEETIROQDEVGDRDRRPEQKKAKGUKGITLLM
QTLNTLSTPEKTLAALCKYAEILLESHNSOKMKLQKXOSOLVQEKHOLRGEHSA
VLARKLESUCRLRNLNRSLKEEGVQRAREEKEKVTSHPQVTLNDILOMQEHN
ERNSKLUQNMELAEARLKLIEQYEURREHIDKVFHKDQQQLVDAKUQAQEMLIKE
AEPHQREKDFLKEAVESQRMCLMKMLKQHLKQALYTERKPEFQNTLSKSEVF
TTFQGEKMTKKIKLEKETIYWRKWSSEKNALEMAEKTVRKLEGLQVKIQIR
LEKLCALOTERNLKNRVODLGAQGQSLTDSGPERRRPEGPAQAQAPSSPRVTEAPCY
PGAPSTEASGCTGPQEPSTARA"

ORIGIN

Alignment Scores:
Pred. No.: 1,79e-124 Length: 4697
Score: 2702.00 Matches: 530
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-023-523-8 (1-530) x AX239574 (1-4697)

QY	1	LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln	20
Db	3	AAAAGCAGCCAGGACAAACCGAAGCAGGACCCGAGGAGCCAGGAGCGCCAGCAG	62
QY	21	AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu	40
Db	63	GCGGCTCTCGAGTAGAAGCAGAAAGTCCCGCGCAGCAGCCAGGCTCTCTCGAAGCCGGAG	122
QY	41	GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer	60
Db	123	GGGGCTCAAGCCAGAACGGCTCAGCTTGGGGCCCTTCGTGATGTCTCTGAGGAGCTCAGC	182
QY	61	ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGly	80
Db	183	CGCCAACTGGAAGACATACCTGAGCACATACCTGTGTGACAATAACCGAGGGGGCCCCGGC	242
QY	81	GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr	100
Db	243	GAGGATGGGCAACAGGGTGAGCCGGCTGAACCCGAAGATGCAGAGAAGTCCCGGACCTAT	302
QY	101	ValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGluLysGluProSer	120
Db	303	GTGCAAGGAATGGGGAGCCCTGGAACCAACTCCAGTAGCTCTATGAGAGAGAAGAACCCCTCC	362
QY	121	LysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHis	140
Db	363	AAGGGGATCCAAACACAGAAAGATCCGGCAGAGTGCAGAGGTCCGAGACCCGAGACCAT	422
QY	141	ArgArgProGlnGluLysLysAlaIAscGlyLeuGlyLysGluIleThrLeuLeuMet	160
Db	423	CGAAGGCCACAGGAGAAGAAAAAGCAAGGGTTTGGGGAGGAGATCACGTTGCTCATG	482
QY	161	GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysTyr	180
Db	483	CAGACATGAAATCTCTGAGTACCCACAGAGGAGAGCTGGCTGCTCTCTGCAAGAGTAT	542
QY	181	AlaGluLeuGluGluHuHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLys	200
Db	543	GCTGAATGTTGGAGGAGCAGCCGAATTTCAGAAAGCAGATGAAGCTCTTACAGAAAAAG	602


```

Db      1563 GGGGCTCAGAGCCCACTCCGCGAGGGCC 1592
|||||
RESULT 6
AR409338
LOCUS   Sequence 46 from patent US 6632923.
DEFINITION
ACCESSION AR409338
VERSION   AR409338.1 GI:40160126
KEYWORDS
SOURCE   Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1638)
AUTHORS  Lees A.M., Lees R.S., Law S.W. and Arjona A.A.
TITLE    Low density lipoprotein binding proteins and their use in
          diagnosing and treating atherosclerosis
JOURNAL  Patent: US 6632923-A 46 14-OCT-2003;
FEATURES
         Location/Qualifiers
         1..1638
            /organism="unknown"
            /mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.:      1.64e-124      Length:      1638
Score:          2693.00        Matches:    529
Percent Similarity: 99.81%    Conservative: 0
Best Local Similarity: 99.81% Mismatches: 1
Query Match:    99.67%        Indels:     0
DB:             6             Gaps:       0

US-10-023-523-8 (1-530) x AR409338 (1-1638)

QY      1 LysSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 20
Db      49 AANAAGCAGCCAGACACACCGGAAGCAGGACCCGAGGAGGCCAGGAGCGGCCAGCAG 108
QY      21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40
Db      109 GCGGCTCTGCTAGTAGAAGCAGAGAGGTCGCGGAGCAGGAGGCTCTCGGAAGCCGAG 168
QY      41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluLeuSer 60
Db      169 GGTGCTCAAGCAGCAAGCGCTCAGTCTGGGGCCCTTCGTGATGTCTCTGAGGAGCTGAG 228
QY      61 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGly 80
Db      229 CGCCAACTGGAGACATCTGAGCACATCTGTCTGGACATAACACAGGGGGGCCCCGGC 288
QY      81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
Db      289 GAGGATGGGGCACAGGGTGAGCGGCTGAACCCGAAGATGCAGAGAAAGTCCCGACCTAT 348
QY      101 ValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGlyLysGluProSer 120
Db      349 GTGGCAGGAATGGGAGCCCTGAACCAACTCCAGTAGTCAATGGAGAGGAACCTCTCC 408
QY      121 LysGlyAspProAsnThrGluLeuArgGlnSerAspGluValGlyAspArgAspHis 140
Db      409 AAGGGGATCCAAACACAGAGAGATCCGGCAGAGTGACGAGTGCCGAGACCGAGACCAT 468
QY      141 ArgArgProGlnGlyLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMet 160
Db      469 CGAAGGCCACAGGAGAGAAAAGCCAAAGGGTTGGGTAAAGGAGATCACGTTCTGATG 528
QY      161 GlnThrLeuAsnThrLeuSerThrProGluLysLysLeuAlaLeuCysLysLysTyr 180
Db      529 CAGACATTGANTACTCTGAGTACCCAGAGAGAGCTGGCTGCTCTGTGCAAGAGTAT 588
QY      181 AlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLys 200
Db      589 GCTGAACCTGCTGGAGGAGCACCGGAATTCACAGAGCAGAGATGAAGCTCTACAGAAAAG 648

```

RESULT 7

AX239603

LOCUS

DEFINITION

AX239603 1638 bp DNA linear PAT 26-SEP-2001
 Sequence 46 from Patent WO0164874.

521 GlyProGlnGluProThrSerAlaArgAla 530
 1609 GGGCCTCAAGAGCCCACTCCGCGAGGGCC 1638

QY 201 GlnSerGlnLeuValGlnGlnLysAspHisLeuArgGlyGluHisSerLysAlaValLeu 220
 Db 649 CAGAGCCAGCTGTGCAAGAGAGGACACACCTTCGCGGTGACACAGCAGGCCCTCTCTG 708
 QY 221 AlaArgSerLysLeuGlnSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLys 240
 Db 709 GCCCGCAGCAAGCTTGAGAGCCTATGCCGTGAGCTGCAGCGGCACAAACCGCTCCCTCAAG 768
 QY 241 GluGluGlyValGlnArgAlaArgGlnGluGluGlyLysArgLysGluValThrSerHis 260
 Db 769 GAAGAAGGTGTGACGGGCCCCGGAGAGGAGGAGGAGGCAAGCGAGAGGTACCTCGCAC 828
 QY 261 PheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSer 280
 Db 829 TTCCAGTGACACTGATGACATTTCAGCTGCAGATGGAACACACACATAGAGCGCACTCC 888
 QY 281 LysLeuArgGlnGlnAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGlnTyr 300
 Db 889 AAGCTCGGCCAAGAGAACATGGAGCTGGCTGAGAGGCTCAAGAAGCTGATTGAGCAGTAT 948
 QY 301 GluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnGlnLeu 320
 Db 949 GAGCTGCGGAGAGGATATCCAAAGTCTTCAACACACAGAGACTACACAGCAGCTG 1008
 QY 321 ValAspAlaLysLeuGlnGlnAlaGlnMetLeuLysGluAlaGluArgHisGln 340
 Db 1009 GTGGATGCCAAGCTCCAGCAGGCCAGAGATGCTAAAGAGGAGGAGAGGCGGCCAG 1068
 QY 341 ArgGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMet 360
 Db 1069 CGGGAGAGGATTTCTCTGAAAGAGGAGTAGAGTCCAGAGGATGTGAGCTGATG 1128
 QY 361 LysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGlu 380
 Db 1129 AAGCAGCAAGAGAGACCCACCTGAAGCAACAGCTTGCCCTATACACAGAGAAGTTGAGGAG 1188
 QY 381 PheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGlnMetGlu 400
 Db 1189 TTCAGAACACACTTTTCAAAAGACGAGGAGTATTCACCACTTCAAGCAGAGATGGAA 1248
 QY 401 LysMetThrLysLysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArgTyr 420
 Db 1249 AGATGACTAAGAGATCAAGAGCTGGAGAAAGAAACCACTGATCCGCTCCCGGTGG 1308
 QY 421 GluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGlu 440
 Db 1309 GAGAGCAGCAACAAGGCCCTGCTTGAGATGGCTGAGGAGAAAAACAGTCCGGGATAAGAA 1368
 QY 441 LeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThr 460
 Db 1369 CTGGAGGGCTCGAGGTAAAAATCCAACGGCTGGAGAACTGTGCGGCACTGCAGACA 1428
 QY 461 GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeu 480
 Db 1429 GAGCGCAATGACCTGAACAAGAGGGTACAGGACCTGAGTGTGTTGGCCAGGGCTCCCTC 1488
 QY 481 ThrAspSerGlyProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerPro 500
 Db 1489 ACTGACAGTGGCCCTGAGAGGAGGCGAGGGGCTGGGGCTCAAGACCCAGCTCCCCC 1548
 QY 501 ArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThr 520
 Db 1549 AGGTCACAGAGAGCGCTTGTCTACCCAGGAGCACCGAGGACAGAGCATCAGGCCAGCT 1608


```

QY 501 ArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThr 520
Db 1549 AGGGTCACAGAGCGCCCTTGTACCCAGAGCAGCAGCAGAGCAGAGCAGATCAGGCCAGACT 1608

QY 521 GlyProGlnGluProThrSerAlaArgAla 530
Db 1609 GGGCTCAAGAGCCACCTCCGCCAGGSCC 1638

RESULT 8
AF516206
LOCUS Homo sapiens taxilin mRNA, complete cds.
DEFINITION AF516206
ACCESSION AF516206.1 GI:28394223
VERSION AF516206.1
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1993)
AUTHORS Nogami, S., Satoh, S., Nakano, M., Shimizu, H., Fukushima, H.,
Maruyama, A., Terano, A. and Shirataki, H.
TITLE Taxilin; a novel syntaxin-binding protein that is involved in
Ca2+-dependent exocytosis in neuroendocrine cells
JOURNAL Genes Cells 8 (1), 17-28 (2003)
MEDLINE 22448667
PubMed 12558796
REFERENCE 2 (bases 1 to 1993)
AUTHORS Shirataki, H., Nogami, S. and Satoh, S.
JOURNAL Direct Submission
TITLE Submitted (29-MAY-2002) Division of Molecular and Cell Biology,
Institute for Medical Sciences, Dokkyo University School of
Medicine, 880 Kitakobayashi, Mibu-machi, Tochigi 321-0293, Japan
LOCATION/Qualifiers
FEATURES
source
1. .1993
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="1"
/map="lp36.11-p34.2"
/tissue_type="brain"
79. .1719
/codon_start=1
/product="taxilin"
/protein_id="AA042465.1"
/db_xref="GI:28394224"
/translation="MKQDKNGAKOSNPKSPQPEAGPEGAQERPSQAAPVAE
GPSSQAPRPEGAQTAQSGALRVSSELSQLEDILSTYCVDNNGGPGDGAQ
EPAPDAKSRITYVARGPEPTPVVNGSEKSPDPTTEIRQSDVEGDRDRPQ
EKKAQGLGKEITLLMOTLNTLSTPEEKLAALCKYAEILLEHRNSOKQMLQKQK
QVCEKDLRGEHAKVLARSKLESLCRELQHRNLSKEGVORAREEEKREXVTS
FOVTLNIOIQMECHNRNSKLRQENMELAEKLLFOVELREHIDLVKFKDLQ
QLVDAKLQQAQEMLEKEERHOREKDFLKEAVESQRMBLQKQETHLQKQALYTE
KFEFQNTLSKSSSVFTTFKQEMEKTKIKLLEKETTMYRSWSSNKALLEMAEK
TVRDKLEGLQVKIQRKLCRALQTERNDLNKRVDLSAGGGGSLTDSGPERRPEP
GAQAPSSPRVTEAPCPGAPSTASGQTGPQETSARA"
ORIGIN
Alignment Scores:
Pred. No.: 2,02e-124 Length: 1993
Score: 2693.00 Matches: 529
Percent Similarity: 99.81% Conservative: 0
Best Local Similarity: 99.81% Mismatches: 1
Query Match: 99.67% Indels: 0
DB: Gaps: 0

US-10-023-523-8 (1-530) x AF516206 (1-1993)

QY 1 LysSerProGlyGlnProGluAlaGlyProGlyAlaGlnGluArgProSerGln 20
Db 127 AAAAGCAGCCAGACAAACCGAGCAGGAGCCCGAGGAGCCAGGAGCGCCAGCAG 186

QY 21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40

```

```

187 GCGGCTCTCTGAGTAGAAGCAGAAAGTCCCGGCGAGCAGCAGGCTCTCGAAGCGGAG 246
QY 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 60
247 GGGGCTCAAGCCAGAAACGGCTCAGCTGGGGCCCTTCGTGATGTCCTCAGAGAGCTGAGC 306
QY 61 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGly 80
307 CGCCAACTGGAGACATACCTGAGACATATCTGTGGACAAATAACAGGGGGGCGCCCGGC 366
QY 81 GluAspGlyAlaGlnGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
367 GAGGATGGGGCAGAGGCTGAGCGGCTGAACCGGAGATGCGAGAGTCCCGGACCTAT 426
QY 101 ValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGluLysGluProSer 120
427 GTGGCAAGGATGGGGAGGCTGAACCACTCCAGTAGTCAATGGAGAGAGGACCTCTC 486
QY 121 LysGlyAspProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAspHis 140
487 AAGGGGGATCCAAACACAGAGAGATCCGGCAGAGTGACGAGTCCGAGACCGGAGCAT 546
QY 141 ArgArgProGlnGluLysLysValAlaLysGlyLeuGlyLysGluLeuLeuMet 160
547 CGAAGGCCACAGGAGAGAAAGCCAAAGGTTTGGGAGAGAGATCAGTTGCTGATG 606
QY 161 GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysTyr 180
607 CAGACATTGAATACTCTGAGTACCCAGAGGAGAGTGGCTGCTCTGTGCAAGAGATAT 666
QY 181 AlaGluLeuGluGluLysHisArgAsnSerGlnLysGlnMetLysLeuGlnLysLys 200
667 GCTGAATCTCTGAGGAGCAGCCGGAATCACAGAGCAGATGAAGCTCTCAGAAAAAG 726
QY 201 GlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGlyLysLysLysAlaValLeu 220
727 CAGAGCCAGCTGGTGCACAGAGAGGACCACTCGCGGCTGAGCAGCAGGCGCTCTG 786
QY 221 AlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLys 240
787 GCGCGCAGCAAGCTTGAGAGCCTATGCGCTGAGCTGCGCGGACACCAACGCTCTCTCAG 846
QY 241 GluGluGlyValGlnAlaArgAlaArgGluGluGluLysArgLysGluValThrSerHis 260
847 GAAGAGGCTGTCGACCGCGCGCGCGGAGGAGAGAGCGGACGAGGAGTGACCTCGCAC 906
QY 261 PheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSer 280
907 TTCAGGTCACACTGAATGACATTCAGCTGCGAGATGGAACAGACACAATGAGCGCAACTCC 966
QY 281 LysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuLeuGlnTyr 300
967 AAGCTCGGCCAGAGAACATGAGCTGGCTGAGAGGCTCAAGAAGCTGATTGAGCAGTAT 1026
QY 301 GluLeuArgGluGluHisLysLysValPheLysHisLysAspLeuGlnGlnLeu 320
1027 GAGCTCGCGGAGGAGCATATCGACAAAGTCTTCAAACACAGAGGACCTACACAGCAGCTG 1086
QY 321 ValAspAlaLysLeuGlnGlnAlaGlnLeuMetLeuLysGluAlaGluArgHisGln 340
1087 GTGGATGCCAGCTCCAGAGGCGCCAGGAGATGCTAAAGGAGGAGCAGAGAGCGGACCAAG 1146
QY 341 ArgGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMet 360
1147 CGGGAGAGGATTTCTCTCTGAAGAGGAGTAGAGTCCAGAGGATGTTGTCAGCTGATG 1206
QY 361 LysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGlu 380
1207 AAGCAGCAGAGAGCAGCCACCTGAAGCAACAGCTTGGCTTATACACAGAGAGTTTGAGGAG 1266
QY 381 PheGlnAsnThrLeuSerLysSerGluValPheThrThrPheLysGlnGluMetGlu 400

```

1267	DB	TTCAGAACACACTTTTCCAAAAGCAGCAGGATTTCCACCACTTCAAGCAGGAGATGGAA	1326
401	QY	LysMetThrLysLysLysLysLeuGluLysGluThrThrMetTyrArgSerArgTTP	420
1327	DB	AAGTGAAGTAAAGAGTCAAGAGCTGGAGAAAGAACCAACATGTACCGGTCCCGGTGG	1386
421	QY	GlusSerSerAsnLysAlaLeuLeuGluMetAlaGluLysThrValArgAspLysGlu	440
1387	DB	GAGAGCAGCAACAAAGCCCTGTCTTCAGATGGCTGAGGAGAAACAGTCCGGGATAAGAA	1446
441	QY	LeuCluLysLeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThr	460
1447	DB	CTGGAGGGCTGCAGGTAAATTCACCGCTGGAGAGACTGTCCCGGGACATGCAGACA	1506
461	QY	GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeu	480
1507	DB	GAGCGCAATCACCTGAACAGAGGGTACAGGACCTGAGTGTGGTGGCCAGGGGTCCCTC	1566
481	QY	ThrAspSerGlyProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerPro	500
1567	DB	ACTCACAGTGGCCCTCAGAGGAGGCGAGAGGGGCTGGGGCTCAAGCACCACTCCCC	1626
501	QY	ArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThr	520
1627	DB	AGGTCACAGAGCGCTTGCTACCCAGGAGCACCGAGCAGACAGAAAGCATCAGGCCAGCT	1686
521	QY	GlyProGlnGluProThrSerAlaArgAla	530
1687	DB	GGGCTCAAGAGCCACCTCCGCCAGGGGCC	1716
RESULT 9			
HSN803946	LOCUS	4880 bp mRNA linear PRI 17-JUN-2003	
DEFINITION Homo sapiens mRNA; cDNA DKFP451I0918 (from clone DKFP451I0918); complete cds.			
AL832636	ACCESSION		
AL832636.2	VERSION	GI:30268324	
KEYWORDS Homo sapiens (human)			
SOURCE Homo sapiens			
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE 1 (bases 1 to 4880)			
AUTHORS Ansoerge,W., Krieger,S., Regiert,T., Rittmueller,C., Schwager,B., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.			
TITLE Direct Submission			
JOURNAL Submitted (17-JUN-2003) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY			
COMMENT On Apr 30, 2003 this sequence version replaced gi:21733212. Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by EMBL (European Molecular Biology Laboratories, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.			
This clone (DKFP451I0918) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 5, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at http://mips.gsf.de/proj/cDNA/.			
FEATURES			
source			
1..4880			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="RZPD:DKFP451I0918"			
/db_xref="taxon:9606"			
/map="lp34.2-36.11"			
/clone="DKFP451I0918"			
/tissue_type="human spinal cord"			
/clone_lib="451 (synonym: hicc1). Vector pSport1; host DH10B; sites NotI + SalI"			
/dev_stage="adult"			
1..4880			
gene			

/gene="DKFZp451I0918"
 125. .1765
 /gene="DKFZp451I0918"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="CAD8951.1"
 /db_xref="GI:30268325"
 /translation="WKNDKXNGAQQSNKSSPGOPEAGPEGAQRPSQAAAPAVEAE
 GPSSQAFKPGGAQRQAQGAARDVSEELSLRQLEDILSTYCVDNQQGPPEDGGAQ
 EPAEPDAKSTYVARGNEPEPTFVNGEKSPGQDNTTEIRQSDVGDRRHRPQ
 EKKAAGKGEITLLMQTLNTLSTPEKLAALCKKYAELLEHNQKMKLQKKQS
 QLVQEKHLRGHSHXAVLRAKSLSLRELQRHNRSLKESGVQARAEERKEKVTSH
 FQVTLIDLOQWENENSLROBNVELARLKKLLEQVELREZHIIDRVFKHKLDIQQ
 QLVDAKLQAOAEWLKEAREHQREKDFILKEAVESQSMCELMKQOETHLQQLALAYTE
 KFEQFNTLSKSSSEVFTTEGEMEKWKIKLLEKETTMYSRSESSNKALLEMAEEK
 TVRDELGLQVTKRLKRLKALQTERNDLNRKVODLSAGGQSLTDSGPPRRPEGP
 GQAQSSPRVTEAPCPGAPSTAEALQTERNDLNRKVODLSAGGQSLTDSGPPRRPEGP
 4833. .4838
 /gene="DKFZp451I0918"
 4852
 /gene="DKFZp451I0918"
 ORIGIN
 Alignment Scores:
 Pred. No.: 8.15e-124 Length: 4880
 Score: 2689.00 Matches: 528
 Percent Similarity: 99.81% Conservative: 1
 Best Local Similarity: 99.62% Mismatches: 1
 Query Match: 99.52% Indels: 0
 Ds: 9 Gaps: 0
 US-10-023-523-8 (1-530) x HSM803946 (1-4880)
 1 LysSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 20
 173 AMAAGCAGCCAGGACACCGAGCAGGACCCGAGGAGCCAGGAGCGCCAGCCAGCAG 232
 21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40
 233 GCGGCTCTCTGACGTAGAAGCAGAAAGTCCCGCAGCAGCCAGGCTCTCTCGAAGCCGAG 292
 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 60
 293 GGGGCTCAAGCCAGACGGCTCAGTCTGGGGCCCTTCGTGATGCTCTGAGGAGCTGAGC 352
 61 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGly 80
 353 CGCCAACTGGAAGACATACTGAGCACATACCTGTGTGGCAATAACCGAGGGGGCCCGCC 412
 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
 413 GAGATGGGGCACAAGGTCAGCGGCTGACCCGAGATGAGAGAGAGTCCCGGACCTAT 472
 101 ValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGluLysGluProSer 120
 473 GTGGCAAGGAATGGGAGGCTCAACCACTCAGTAGTCAATGGAGAGAGAACCCCTCC 532
 121 LysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHis 140
 533 AAGGGGATCCAAACACAGAGAGATCCGGCAGATGACGAGTGCAGAGTCCGAGACCGAGCAT 592
 141 ArgArgProGlnGluLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMet 160
 593 CGAAGGCCACAGAGAGAAGAAAGCAAGGGTTGGGGAAGGAGATCACTTGCTGATG 652
 161 GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuIleAlaLeuCysLysLysTyr 180
 653 CAGACATTTGAATTACTCTGAGTACCCAGAGGAGAGCTGGGCTGCTCTGTGCAAGAGTAT 712
 181 AlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLys 200
 713 GCTGAATCTGTCAGAGACACCGGAATTCACAGAGCAGATGAAGCTCTCTAGAGAAAG 772

gene

21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40
 211 CGGCTCTCTGAGTAGAGAGAGAGAGTCCCGGAGCAGCAGGCTCTCGGAGCGGAG 270
 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 60
 271 GGGGCTCAAGCCAGCAAGCGCTCAGTCTGGGGCCCTTGGTGTGCTCTGAGGAGCTGAGC 330
 61 ArgGlnLeuGluAspLeuSerThrTyrcysValAspAsnGlnGlyGlyProGly 80
 331 CCCCACTGGAGAGATACCTAGCAGCATACCTGTGGACATACCAAGGGGGGCCCCGGC 390
 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTy 100
 391 GAGGATGGGACACAGGCTGAGCGGCTGAACCGAGAGATGCGAGAGAGTCCCGGACCTAT 450
 101 ValAlaArgAsnGlyGluProGluProThrProValValTyrcysGluLysGluProSer 120
 451 GTGGCAAGGAATGGGGAGCTGAACCAACTCCAGTAGTCAATGGAGAGAGAACCTCTCC 510
 121 LysGlyAspProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAspHis 140
 511 AAGGGGGATCCAAACACAGAGAGATCCGGCAGAGTACGAGGTCGGAGAGACCGAGCAT 570
 141 ArgArgProGlnGluLysLysLysLysLysGlyLeuGlyLysGluLeuThrLeuMet 160
 571 CGAAGGCCACAGGAT 630
 161 GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaAlaLeuCysLysLysTy 180
 631 CAGACATTGAATCTCTGAGTACCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 690
 181 AlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLysLysLys 200
 691 GCTGAACCTGCTGAGGAT 750
 201 GlnSerGlnLeuValGlnLysLysLysLysLysLysLysLysLysLysLysLysLys 220
 751 CAGAGCCAGCTGGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 810
 221 AlaArgSerLysLeuGluSerLysLysLysLysLysLysLysLysLysLysLysLys 240
 811 GCCCGCAGCAAGCTTTCAGAGCTATGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 870
 241 GluGluGlyValGlnArgAlaArgGluGluGluGluGluGluGluGluGluGluGlu 260
 871 GAAGAAGGTGTGACGGGGCCCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 930
 261 PheGlnValThrLeuAsnAspLeuGlnLeuGlnMetGluGlnHisAsnGluArgAsnSer 280
 931 TTCAGGTGACACTGATGACATTCAGCTGACAGTGGAGACAGACACAAATGAGCGCACTCC 990
 281 LysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLysLysLysLysLys 300
 991 AAGCTCGCCCAAGAGAACTGAGCTGGCTGAGAGGCTCAAGAGCTGATTGAGCAGTAT 1050
 301 GluLeuArgGluGluHisLysLysLysLysLysLysLysLysLysLysLysLysLys 320
 1051 GAGCTCGCGGAGAGAGATATGCAAAAGTCTTCAACACACAGAGACTACACAGCAGCTG 1110
 321 ValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluGluArgHisGln 340
 1111 GTGGATGCCAAGCTCCAGCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1170
 341 ArgGluLysAspPheLeuLysLysLysLysLysLysLysLysLysLysLysLysLys 360
 1171 CGGGAGAGAGATTTCTCTCTGAAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAT 1230
 361 LysGlnGlnGluThrHisLysLysLysLysLysLysLysLysLysLysLysLysLys 380
 1231 AAGCAGCAAGAT 1290
 381 PheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGlu 400

1291 TTCCAGAACACACTTTCCAAAAAGCAGGAGGATTTCACACATTCACAGGAGATGAA 1350
 401 LysMetThrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 420
 1351 AAGATGACTAAGAGATCAAGAGCTGGAGAAAGAAACACCATGTACCGGTCCCGGTGG 1410
 421 GluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGlu 440
 1411 GAGAGCAGCAACAAGCCCTGCTTCAGATGGCTGAGGAGAAACAGTCCGGGATAAAGAA 1470
 441 LeuGluGlyLeuGlnValLysLysLysLysLysLysLysLysLysLysLysLysLys 460
 1471 CTGAGGGGCTGAGAGTAAATCAACGGCTGGAGAGAGTGTGCGGGCACTGCAGACA 1530
 461 GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeu 480
 1531 GAGCGCAATGACCTGAACAAAGAGGGGTACAGGACCTGAGTGTGGTGGCCAGGGCTCCCTC 1590
 481 ThrAspSerGlyProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerPro 500
 1591 ACTGACAGTGGCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1650
 501 ArgValThrGluAlaProCysTyrcysProGlyAlaProSerThrGluAlaSerGlyGlnThr 520
 1651 AGGTCACAGAGAGCGCTTGTACCCAGGAGCAGCAGGAGCAGAGAGATCAGGCCAGACT 1710
 521 GlyProGlnGluProThrSerAlaArgAla 530
 1711 GGGGCTCAAGAGCCCACTCCCGCAGGGCC 1740

RESULT 11
 HSM803947 5023 bp mRNA linear PRI 17-JUN-2003
 LOCUS Homo sapiens mRNA; cDNA DKFZp451J0118 (from clone DKFZp451J0118).
 DEFINITION AL832637
 ACCESSION AL832637.2 GI:30268326
 VERSION
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 5023)
 AUTHORS Ansorge,W., Krieger,S., Regiert,T., Rittmüller,C., Schwager,B.,
 Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
 Wiemann,S.
 TITLE Direct Submission
 JOURNAL Submitted (17-JUN-2003) MIPS, Ingolstaedter Landstr.1, D-85764
 Neuherberg, GERMANY
 COMMENT On Apr 30, 2003 this sequence version replaced gi:21733213.
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by EMBL (European Molecular Biology Laboratories,
 Heidelberg/Germany) within the cDNA sequencing consortium of the
 German Genome Project.
 This clone (DKFZp451J0118) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clonerzpd.de Further
 information about the clone and the sequencing project is available
 at http://mips.gsf.de/proj/cDNA/.

FEATURES
 source
 1. 5023
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="RZPD:DKFZp451J0118"
 /db_xref="taxon:9606"
 /map="1p34.2-36.11"
 /clone="DKFZp451J0118"
 /tissue_type="human spinal cord"
 /clone_lib="451 (synonym: hlcc1). Vector pSport1; host
 DH10B; sites NotI + SalI"
 /dev_stage="adult"
 1. 5023

gene

KEYWORDS
SOURCE
ORGANISM

Unknown.
Unknown.
Unclassified.

REFERENCE

1 (bases 1 to 4722)
Lees, A.M., Lees, R.S., Law, S.W. and Arjona, A.A.
Low density lipoprotein binding proteins and their use in
diagnosing and treating atherosclerosis

JOURNAL
Patent: US 6605588-A 14 12-AUG-2003;

FEATURES

Location/Qualifiers
1..4722
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 8 16e-113 Length: 4722
Score: 2465.50 Matches: 492
Percent Similarity: 92.99% Conservative: 12
Best Local Similarity: 90.77% Mismatches: 25
Query Match: 91.25% Indels: 13
DB: 6 Gaps: 3

US-10-023-523-8 (1-530) x AR374685 (1-4722)

QY 1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 20
DB 109 AAAAGCAGCCCGGACAGCGGAGCAGGAGCGGAGCGGAGCGGCGGCGGCGG 168
QY 21 AlaAlaProAlaValGluAlaGlyProGlySerSerGlnAlaProArgLysProGlu 40
DB 169 CCGGCCCCCGCGAGACCGGAGGT---GCCAGCAGCAGGCTCCCGGAGGCGGAG 225
QY 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluLeuSer 60
DB 226 GGGGCTCAAGCAAACTGCTCAGCGTGGGGGCTCTGTCATGCTCTCTCAGGAGCTGAGC 285
QY 61 ArgGlnLeuGluAspIleLeuSerThrTyrcysValAspAsnAsnGlnGlyProGly 80
DB 286 CGCAGTTGGAACATATCTAGTATCATCTGTGTGACAAACACGAGGGGCGGCGGT 345
QY 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
DB 346 GAGATGGGGTCCAGGGTGAAGCCCTGAACTGAACTGAACTGAACTGAACTGAACT 405
QY 101 ValAlaArgAsnGlyGluProGluPro---ThrProValValTyrcysGlyGluLysGluPro 119
DB 406 GTGGCAAGGAATGGGAGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 465
QY 120 SerLysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAsp 139
DB 466 TCCAGGCGAGCGCGGCGGCGGAGGATCCGAGCGAGCGATGAGTCCGAGACCGAGAC 525
QY 140 HisArgArgProGlnGluLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeu 159
DB 526 CACCGGAGGCCACAGGAAAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 585
QY 160 MetGlnThrLeuAsnThrLeuSerThrProGluLysLeuAlaAlaLeuCysLysLys 179
DB 586 ATGCAGACATGACACGCTGAGCACCACCGAGGAGAGGCTGCGGCTCTGTGCAAGAG 645
QY 180 TyrAlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLys 199
DB 646 TATGGCGAATGCTCGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 705
QY 200 LysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaVal 219
DB 706 AAGCAGAGCCAGCTGGTGGAGGAGAGGACCTCGCTGGCGGAGCAGCAGCAGGCGCATC 765
QY 220 LeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeu 239
DB 766 CTGGCCCCGACGAGCTCGAGAGGCTGTGCGGAGGCTGTGCGGAGGCTGCGGCGGCGG 825

QY 240 LysGluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSer 259
DB 826 AAGGAGAGAGGTGTGCAGCGAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 885
QY 260 HisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsn 279
DB 886 CACTTCAGATGAGCTCAACGACATTCAGTCGAGTGGAGCAGCAGCAGCAGCAGCAG 945
QY 280 SerLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLysLysLeuLeuGluGln 299
DB 946 TCCAAGCTGGCCAGGAGAACATGGAGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1005
QY 300 TyrGluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGln 319
DB 1006 TACGAGCTGGCAGAGAGACATCGCAAAAGTCTTCAACACAGAGATCTGCAGCAGCAG 1065
QY 320 LeuValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHis 339
DB 1066 CTGGTGAGCGCCAGAGCTCCAGCGGCCAGAGATGCTGAAGAGGCGAGGAGCGGCGAC 1125
QY 340 GlnArgGluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeu 359
DB 1126 CAGCGGAGAGAGACTTTCTCTGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1185
QY 360 MetLysGlnGlnGluThrHisLysLysGlnGlnLeuAlaLeuTyrcysGluLysPheGlu 379
DB 1186 ATGAGCAACAGGAGAGCCACCTGAAGCAGCAGCTTCCCTATACACAGAGAGTTTGAG 1245
QY 380 GluPheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMet 399
DB 1246 GAGTTCCAGAACACTCTTTCCAAAGAGCAGGAGGTGTTCACCATTCACAGAGGAATG 1305
QY 400 GluLysMetThrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 419
DB 1306 GAAAGATGACAAAGAGATCAAGAGCTGGAGAAAGAGAGAGAGAGAGAGAGAGAGAG 1365
QY 420 TrpGluSerSerAsnLysAlaLeuLeuGluMetAlaGluLysThrValArgAspLys 439
DB 1366 TGGGAGAGCACAACAGGCGCTCTGAGATGCTGAGGAGAGAGAGAGAGAGAGAGAG 1425
QY 440 GluLeuGluGlyLeuGlnValLysIleGlnArgGluGluLysLeuCysArgAlaLeuGln 459
DB 1426 GAGCTGGAGAGCGCTCGAGGTGAAATCCAGCGCTGGAGAGAGCTGTGCGGCGGCTGCAG 1485
QY 460 ThrGluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySer 479
DB 1486 ACAGAGCGCATGACCTGAACAAGAGGTGCAGGACCTGAGTGGCGGTGGCGGCGGCGG 1545
QY 480 LeuThrAspSerGlyProGluArgPro----- 489
DB 1546 GTCTCCGACAGCGGTCTCTGAGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1605
QY 490 ---GluGlyProGlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyr 508
DB 1606 GTCGAGGCGCGCGGCGCTCAAGTACCCACTCTCCAGGGCCACAGAGCGCTCTCTGCTGC 1665
QY 509 ProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAla 528
DB 1666 GCAGGTGCACCGCAGCAGAGAGGATCAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1725
QY 529 ArgAla 530
DB 1726 ACTGCC 1731

RESULT 14

AR409320 4722 bp DNA linear PAT 18-DEC-2003
LOCUS AR409320
DEFINITION Sequence 14 from patent US 6632923.
ACCESSION AR409320
VERSION AR409320.1 GI:40160108
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.

Unclassified.

1 (bases 1 to 4722)

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.
Low density lipoprotein binding proteins and their use in
diagnosing and treating atherosclerosis
Patent: US 632923-A 14 14-OCT-2003;
Location/Qualifiers
1. - 4722
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 8.16e-113 Length: 4722
Score: 2465.50 Matches: 492
Percent Similarity: 92.99% Conservative: 12
Best Local Similarity: 90.77% Mismatches: 25
Query Match: 91.25% Indels: 13
DB: 6 Gaps: 3

US-10-023-523-8 (1-530) x AR409320 (1-4722)

QY 1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 20
Db 109 AAAAGACCGCCGGACAGCCGGAAGCAGGAGCGGAGCCCGGGGGCGCCGGCGG 168
QY 21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40
Db 169 CCGGCCCGCCCGAGAGCCGAGGT--GCCAGCAGCCAGGCTCCCGGGAGCCGGAG 225
QY 41 GlyAlaGlnAlaThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 60
Db 226 GGGGCTCAAGCCAAATCTGTCAGCTGGGGCGCTCTGTGATCTCTCAGGAGCTGAGC 285
QY 61 ArgGlnLeuGluAspLeuSerThrTyrcysValAspAsnGlnGlyGlyProGly 80
Db 286 CGCAGTTGAAGACATCTAGTACTACTGTGTGACAAACACCGGGGGCCCGGT 345
QY 81 GluAspGlyAlaGlnGlyProAlaGluProGluAspAlaGluLysSerArgThrTy 100
Db 346 GAGGATGGGCTCAGGGTGAGCCCGCTGAACCTGAAGATGCAGAGAAGTCTCGCGCTAT 405
QY 101 ValAlaArgAsnGlyGluProGluPro--ThrProValValTyrcylGluLysGluPro 119
Db 406 GTGCAAGGAATGGGAGCGCGAGCGCGGACCCCGAGTCAATGCCAGAGAGGAGACC 465
QY 120 SerLysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAsp 139
Db 466 TCCAAGCGAGCGCGGACGGAAGAGATCCGACGAGCGGATGAGTCCGAGACCGAGAC 525
QY 140 HisArgArgProGlnGluLysLysAlaLysGlyLeuGlyGlyGluIleThrLeuLeu 159
Db 526 CACCGAGGCCACAGAAAGAGAGAGGCAAGGGTCTGGGAAGAGAGATCACGCTGCTG 585
QY 160 MetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLys 179
Db 586 ATCCAGACACTGACACGCTGAGCACCCAGAGGAGAGCTGGGGCTCTGTGCAAGAG 645
QY 180 TyrAlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuGlnLys 199
Db 646 TATCGGAACTGCTCAGGAGGACCCGGAATCTCGCAGAGCAGATGAAGTCTGTCAGAG 705
QY 200 LysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaVal 219
Db 706 AGCAGAGCCAGCTGTGTGAGGAGAGGACCACTCGTGGCGAGCAGCAGAGCCATC 765
QY 220 LeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeu 239
Db 766 CTGGCCCGCAGCAAGCTCAGAGAGCTGTGCGGGAGCTGCAGCGGCACACCGCTCGCTC 825
QY 240 LysGluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSer 259
Db 826 AAGGAAGAAGGTGTGAGCGAGCCCGAGAGGAGGAGAGGAGGAGGAGGAGGTGACGTCA 885

QY 260 HisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsn 279
Db 886 CACTTCCAGATGAGCTCAACGACATTCAGCTGAGATGGAGCAGCACACGAGCGCAAC 945
QY 280 SerLysLeuArgGlnGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGluGln 299
Db 946 TCCAGCTGCCCGAGGAGAACATGGAGCTGCCCGCGGCTCAAGAGAGCTGATTGAGCAG 1005
QY 300 TyrGluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnGln 319
Db 1006 TACGAGCTGCGAGAGAGCAGCATCGACAAAGTCTTCAACACAAAGGATCTCGACGAGCAG 1065
QY 320 LeuValAspAlaLysLeuGlnGlnAlaGlnGlnMetLeuLysGluAlaGluGluArgHis 339
Db 1066 CTGTGGAGCCAGCTCCAGAGCCCGAGAGATGCTGAAGAGGAGGAGGAGCGGCAC 1125
QY 340 GlnArgGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeu 359
Db 1126 CAGCGGAGAGAGGACTTCTCTCAAGAGGCGCGTGGAGTCCAGAGAGATGTGGAGCTG 1185
QY 360 MetLysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrcylGluLysPheGlu 379
Db 1186 ATGAAGCAACAGGAGACCCACCTGAAGCAGCAGCTTGCCCTTATACAGAGAAATTTGAG 1245
QY 380 GluPheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMet 399
Db 1246 GAGTTCAGAACACTTTCCTCAAGAGCAGGAGGTTCACCACTTCAACAGGAATG 1305
QY 400 GluLysMetThrLysLysIleLysLysLeuGluLysGluThrMetTyrcylGluArg 419
Db 1306 GAAAGATGACAAAGAGATCAAGAGCTGGAGAAGAGAGACCATGTACCGTTCCCGG 1365
QY 420 TrpGluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLys 439
Db 1366 TGGAGAGCAGCACACAGGCGCTCTTGAGTGGCTGAGGAGAGAAACACTCCGGGACAA 1425
QY 440 GluLeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGln 459
Db 1426 CAGCTGGAAGCGCTGCGAGTGAAATCCAGCGGTGGAGAGAGTGTCCCGGCACTGCAG 1485
QY 460 ThrGluArgAsnAspLeuAsnLysArgValcylAspLeuSerAlaGlyGlyGlnGlySer 479
Db 1486 ACAGAGCGCATGACCTGAACAGAGGTGCGAGACCTGAGTCCCGTGGCCAGGGCCCC 1545
QY 480 LeuThrAspSerGlyProGluArgArgPro----- 489
Db 1546 GTCTCCGACAGCGGTCTGAGCGGAGCCAGAGCGCCGCCACCTCCAAGGAGCAGGGT 1605
QY 490 ---GluGlyProGlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyr 508
Db 1606 GTCAGAGCGCCCGGGCTCAAGTACCACTTCCAAGGGCCACAGACGCTTCTCTGCTGC 1665
QY 509 ProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAla 528
Db 1666 GCAGGTCCACCCAGCAGAGGAGCATCAGGCCAGAGGGCCCCCAGGAGGCCACCACTGCC 1725
QY 529 ArgAla 530
Db 1726 ACTGCC 1731
RESULT 15
LOCUS AX239571 4722 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 14 from Patent WO0164874.
ACCESSION AX239571
VERSION AX239571.1 GI:15797253
KEYWORDS
SOURCE Oryctolagus cuniculus (rabbit)
ORGANISM Oryctolagus cuniculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
REFERENCE 1

AUTHORS

Lees, A.M., Lees, R.S., Law, S.W. and Arjona, A.A.

TITLE Low density liprotein binding proteins and their use in diagnosing

JOURNAL and treating atherosclerosis

Patent: WO 0164874-A 14 07-SEP-2001;

Boston Heart Foundation, Inc. (US)

FEATURES

Location/Qualifiers

source

1..4722

/organism="Oryctolagus cuniculus"

/mol_type="unassigned DNA"

/db_xref="taxon:9986"

61..1734

/note="unnamed protein product"

/codon_start=1

/protein_id="CAC88294.1"

/db_xref="GI:15797254"

/db_xref="RENTREMBL:CAC88294"

/translation="MNQDKNGAAKQNPKSSPOQPEAGAEAGQGRPRPAPAREAE
 GASSQAPPEGAQKAPQALCDVSELSQLEDILSTVCVDNQAQPEDGVQGE
 PEPEDAERAYARNGEPETPVNGEKETSAEPGETEIRTSDEVDHRPQ
 EKKAKGKETTLLMQTLNTSTPEKLAALKKYAELLBEHRNSQKMLLOKQS
 OLVOEKDLRGPHSKAILARSKLBSLCRELQRHNSLKEEGVQVAREEKEKVTSH
 FQMTNLDLOLOMEQHNRNSKLROENMELAEELKLIHQYELRESHIDKVPKHDLQ
 QVLDKLAQQAEMLEKABERHOREKDFLLKEAVESQRCMELMKQETHLKKQALYTE
 KPEEFQNTLSKSEVFTTPKQEMKWKIKLKEKETTYRWSWSSNKLLEWAEK
 TLRLXELGLQVKIQRLKCLRALQTERNDLNKRVQDLSAGQGVPSDSGPRRPEPA
 TTTSKQGVGPGQAVPNSPRATDASCCGAPSTASGQTGPQEPPTATA"

ORIGIN

Alignment Scores:

Pred. No.: 8,166-113 Length: 4722
 Score: 2465.50 Matches: 492
 Percent Similarity: 92.99% Conservative: 12
 Best Local Similarity: 90.77% Mismatches: 25
 Query Match: 91.25% Indels: 13
 DB: 6 Gaps: 3

US-10-023-523-8 (1-530) x AX239571 (1-4722)

QY 1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 20
 DB 109 AAAAGCAGCCCGGAGCAGCGGAGCAGGAGCGGAGGAGCCAGGGGGCGGCGCGG 168
 QY 21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40
 DB 169 CCGGCCCCCGCGGAGAACCGAAGGT---GCCAGCAGCCAGGCTCCCGGGAGGCGGAG 225
 QY 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 60
 DB 226 GGGGCTCAGCCAAACTGCTCAGCTGGGGCGCTCTGTGATGCTCTCAGAGAGCTGAGC 285
 QY 61 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnGlnGlyGlyProGly 80
 DB 286 CGCCAGTTGGAAGACATCTCAGTACATATCTGTGTGACAAACACAGGGGGCGCGGCT 345
 QY 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
 DB 346 GAGGATGGGGTCCAGGGTGAGCCCCCTGAACCTGAGATGCGAGAGTCTCGCGCTAT 405
 QY 101 ValAlaArgAsnGlyGluProGluPro---ThrProValValTyrGlyGluLysGluPro 119
 DB 406 GTGGCAAGAAATGGGAGCGGAGCGGAGCCGCCAGTGTCAATGGCGAGAGGAGACC 465
 QY 120 SerLysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAsp 139
 DB 466 TCCAGGGCAGAGCGGGCAGGAGAGATCCGGAGCGAGATGAGTCCGAGACCGAGAC 525
 QY 140 HisArgArgProGlnGluLysLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeu 159
 DB 526 CACCGGAGCCACAGGAAAGAGAGGACCAAGGGTCTGGGAAAGAGAGATCACGCTGCTG 585
 QY 160 MetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLys 179
 DB 586 ATGCAGACACTGAACACGCTGAGCACCCCGAGGAGAGAGCTGGCGGCTCTGTGCAAGAG 645

QY 180 TyrAlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLys 199
 DB 646 TATCGGAACCTGCTCGAGGAGCACCCGAACTCGCAGAAAGCAGATGAAGCTGCTGCAGAAG 705
 QY 200 LysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaVal 219
 DB 706 AAGCAGAGCCAGCTGGTGCAGGAGAGGACCACTGCTGCGGAGCAGACCAAGGCCATC 765
 QY 220 LeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeu 239
 DB 766 CTGGCCCGCAGCAAGCTCGAGAGCTGTGCGGAGAGTGCAGCGGCACAAACCGCTCGCTC 825
 QY 240 LysGluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSer 259
 DB 826 AAGAAAGAGGTGTGCAGCGAGCCCGAGAGGAGGAGGAAAGCCGACAGAGGTGACGTCA 885
 QY 260 HisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsn 279
 DB 886 CACTTCAGATGACGCTCAACGACATTCAGCTGCAGATGGAGCAGCAGCAACAGCGCAAC 945
 QY 280 SerLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuLeuGln 299
 DB 946 TCCAGCTGCGCCAGGAGAACATGAGCTGCCGAGCGGCTCAAGAGCTGATTGACAG 1005
 QY 300 TyrGluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGln 319
 DB 1006 TACGAGCTCGAGAGAGACATCGACAAAGTCTTCAACACAGAGGATCTGCAGCAGCAG 1065
 QY 320 LeuValAspAlaLysLeuGlnGlnAlaGlnMetLeuLysGluAlaGluArgHis 339
 DB 1066 CTGGTGCAGCGCAAGCTCCAGCAGCGCCAGGAGATGCTGAAGGAGGAGGAGCGGCAC 1125
 QY 340 GlnArgGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeu 359
 DB 1126 CAGCGGAGAGGACTTCTCTGAAGAGGCGCTGGAGTCCCAGAGGATGTGCGAGCTG 1185
 QY 360 MetLysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGlu 379
 DB 1186 ATGAAGCAACAGGAGAGCCCACTGAAGCAGCAGCTGCTCCCTATACACAGAGAAGTTTGAG 1245
 QY 380 GluPheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMet 399
 DB 1246 GAGTTCAGAAACACTCTTCCAAAGCAGCGAGTGTTCACCATTTCAACAGGAATG 1305
 QY 400 GluLysMetThrLysLysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArg 419
 DB 1306 GAAAAGATGACAAAGAAAGATCAAGAACTGGAGAAAGAGACCACTACCTGCTCCCGG 1365
 QY 420 TrpGluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLys 439
 DB 1366 TGGAGAGCAGCAACAAGGGCCCTGCTTGAGATGGCTGAGGAGAAAACACTCCCGGACAA 1425
 QY 440 GluLeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGln 459
 DB 1426 GAGCTGGAAGGCTGCAGGTGAANAATCCAGCGCTGGAGAGCTGTGCCGGGCACCTCAG 1485
 QY 460 ThrGluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySer 479
 DB 1486 ACAGAGCGCATGATGACTGAACAGAGGGTGCAGGACCTGATGCTGCGTGGCGGCCCC 1545
 QY 480 LeuThrAspSerGlyProGluArgPro----- 489
 DB 1546 GTCTCCGACAGCGTCTGTGCGGAGGCGCCAGAGCGCCGCCACCTCCAGGAGCAGGGT 1605
 QY 490 ---GluGlyProGlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyr 508
 DB 1606 GTCGAGGGCCCCGGGCTCAAGTACCACTCTCCAAGGGCCACAGACGCTTCTGTGTCG 1665
 QY 509 ProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAla 528
 DB 1666 GCAGGTGCACCCAGCAGAGGAGCATCAGGCCAGAGGGCCCCAGAGGCCACCCACCTGCC 1725

QY 480 LeuThraspSerGlyProGluArgArgPro----- 489
 Db 1546 GTCTCGACAGCGGTCTCTGAGCGGAGCGCAGCGCCGCCACACCTTCCAGGAGCAGGCT 1605
 QY 490 ---GluGlyProGlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyr 508
 Db 1606 GTGAGGGGCCCCGGGCTCAAGTACCCTCTCCAGGGCCACAGACGCTTCTGCTGTC 1665
 QY 509 ProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAla 528
 Db 1666 GCAGGTGACCCAGCAGCAGGAGCATCAGGCGCAGCAGGCGCCCGCAGGAGCCACCACTGCC 1725
 QY 529 ArgAla 530
 Db 1726 ACTGCC 1731
 RESULT 17
 BC060227
 LOCUS
 DEFINITION Mus musculus cDNA clone MGC:77972 IMAGE:30357541, complete cds.
 ACCESSION BC060227
 VERSION BC060227.1 GI:38614270
 KEYWORDS MGC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 4702)
 Strausberg,R.B., Feingold,E.A., Grouse,L.H., Derge,J.G.,
 Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
 Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
 Diatchenko,L., Marusan,K., Farmer,A.A., Rubin,G.M., Hong,L.,
 Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
 Schetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
 Carninci,P., Prange,C., Rana,S.S., Loquellano,N.A., Peters,G.J.,
 Abramson,R.D., Mullen,S.J., Bosak,S.A., McEwan,P.J.,
 McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
 Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Rulyk,S.W.,
 Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
 Fahey,J., Helton,E., Kettner,M., Madan,A., Rodriguez,S.,
 Sanchez,A., Whiting,N., Madan,A., Young,A.C., Shevchenko,Y.,
 Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
 Butlerfield,Y.S., Krzywinski,M.I., Skalska,U., Smalish,D.E.,
 Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 4702)
 Strausberg,R.
 Direct Submission
 Submitted (20-OCT-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabbs@mail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library Preparation: M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr.
 Thomas L. Casavant
 Web site: <http://genome.uiowa.edu>
 Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
 Bonaldo,M.F., Akabogu,I., Bair,T., Bair,J., Crouch,K., Davis,A.,
 Fishler,K., Keppel,C., Kuabara,T., Lebeck,M., Melo,A., Schaefer,K.,
 Schetz,T., Smith,C., Snir,E., Tack,D., Trout,K., Walters,U.,
 Casavant,T., Soares,M.B.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: Plate: Row: Column: 0.
 Location/Qualifiers
 1. 4702
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /cldones="MGC:77972 IMAGE:30357541"
 /tissue_type="Brain"
 /clone_lib="NIH_BMAP_GHO"
 /lab_host="DH10B"
 /note="vector: pYX-ASC"
 220..1884
 /codon_start=1
 /product="Unknown (protein for MGC:77972)"
 /protein_id="AAH60227.1"
 /db_xref="GI:38614271"
 /translation="MKNDKNGPKAKSNKSGSPGOREAGPEGAHPRPTARGAARE
 GSTSQPKTEGARAKAQAQCALDSELSRQLEDILSTYCVNNOGGPAEGAGGE
 PTEPDEKRTYAARNGPEPGIPVNGEKTSGEPGTEEIRASDEVDRHRPQ
 EKKAKGLGKEITLLMQTLNTSTPEKLAALCKYAELEHRSQKMKLQKQKQ
 QLVQEKDLRGHSHKAVLARSKLESLRELQRHNRSLKEGVORAEERKEKRVTS
 FOYTLNDIQLOMEQHNERNKLRQENMELAEERKLEFLKEAVESORMCELMKQOERLKY
 OLVDKLOQAQAEMLKEAEERHOREKFLKEAVESORMCELMKQOERLKY
 KFEFQNTLSKSEVEFTTPEQEMKTKIKLEKETTMYRSWESNKLLEWAEK
 TVDKLEGLQVKIQRLKLCRALQTERNDLNKRVDLTAGGITDIGSRRRPPATTAS
 KEQGVSPGAPASSPRATDAPCCSGAPSTGTAGTGPGEPTATA"

ORIGIN

Alignment Scores:
 Pred. No.: 1,06e-108 Length: 4702
 Score: 2382.00 Matches: 480
 Percent Similarity: 90.41% Conservative: 10
 Best Local Similarity: 88.56% Mismatches: 36
 Query Match: 88.16% Indels: 16
 DB: 10 Gaps: 4
 US-10-023-523-8 (1-530) x BC060227 (1-4702)

QY 1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 20
 Db 268 AAGGGCAGCCCGGGCAAGCGGAGCAGGAGCGGAGCCATCGAGCAGCAGCAG 327
 QY 21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProGlyProGlu 40
 Db 328 ACAGCTCTGGGGCGGAGCCGAGGT---TCACAGCCAGGCTCTCTGGGAAGCCGAG 384
 QY 41 GlyAlaGlnAlaAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 60
 Db 385 GGGGCTCGAGCTAAAGCAGCTCAGCTGGGGCCCTCTGTGACCTCTCTGAGGAGCTGAGC 444
 QY 61 ArgGluLeuGluAspIleuSerThrTyrcysValAspAsnGlnGlyProGly 80
 Db 445 CGGAGTGTGAAGACATCTGTAGTACATCTGTGTGACCAACAATCAGGAGGCCCGGCT 504
 QY 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
 Db 505 GAGGAGGAGCAGAGGTGAGCCCTGAGCCGAGAGACACGAGAGGTCCCGACCTAT 564
 QY 101 ValAlaArgAsnGlyGluProGluPro---ThrProValValTyrcysGluLysGluPro 119
 Db 565 GCAGCCAGGAATGGGAGGCTGAACAGGCAATTCAGTCTCAACCGCGGAGGAGGAGCC 624
 QY 120 SerLysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAsp 139
 Db 625 TCTAAGGAGAGCCTGCTGAACACAGAGGAGTCCGAGCGAGTGTAGTTCGAGACCGAGAC 684
 QY 140 HisArgArgProGlnGluLysLysLysAlaLysGlyLeuGlyGluLeuLeuLeu 159
 Db 685 CATCGGAGGCCACAGGAGGAAGAAAGCAAGGCTCTAGGAGGAGGAGATCACTCTGCTG 744
 QY 160 MetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLys 179


```

QY 349 GluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLeuLys 368
Db 1311 GAGGAGTAGAGTCCAGAGGATGTGTGAGCTGTATGAGCAGCAGACACCCACCTGAG 1252

QY 369 GlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSer 388
Db 1251 CAACAGCTTGCCCTATACACAGAGAAGTTGAGGAGTTCACAGAACACACTTTCACAAAGC 1192

QY 389 SerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLys 408
Db 1191 AGCGAGGTATTCCACCATTCAGCAGGAGGATGGAAAGATGACTTAAAGATCAAGAAG 1132

QY 409 LeuGluLysGluThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeu 428
Db 1131 CTGGAGAAAGAAACACCATGTACCGTCCCGTGGGAGAGCAGCAACAGGCCCTGCTT 1072

QY 429 GluMetAlaGluGluLysThrValArgAspLysGluLeuGluGluGluValLysLys 448
Db 1071 GAGATGGCTGAGGAGAAACACGTCCGGGATAAGAACTGGAGGGCTCGAGTAAATATC 1012

QY 449 GlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArg 468
Db 1011 CAACGGCTGGAGAGCTGTCCGGGCATCTGCAGACAGAGCGCAATGACTGACACAGAGG 952

QY 469 ValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSerGlyProGluArgArg 488
Db 951 GTACAGGACCTGAGTGTGGTGGCAGGCTCCCTCACTGACAGTGGCCCTGAGAGGAGG 892

QY 489 ProGluGlyProGluValGlnAlaProSerSerProArgValThrGluAlaProCysTyr 508
Db 891 CCAGAGGGCCCTGGGGCTCAAGACCCAGCTCCCGCAGGTCCAGAGAGCGCTTGCTTAC 832

QY 509 ProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAla 528
Db 831 CCAGAGGACCCAGCAGACAGACATCAGCCAGACTGGGCCCTCAAGAGCCACCTCCGCC 772

QY 529 ArgAla 530
Db 771 AGGGCC 766

RESULT 19
HUMHMBGCF/c
LOCUS HUMHMBGCF 1793 bp mRNA linear PRI 25-MAY-1995
DEFINITION Human high molecular weight B cell growth factor mRNA sequence.
ACCESSION L15344
VERSION L15344.1 GI:832913
KEYWORDS B-cell growth factor; interleukin 14.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1793)
Ambrus, J.L., Jr., Pippin, J., Joseph, A., Xu, C., Blumenthal, D.,
Tamayo, A., Claypool, K., McCourt, D., Srikiatchoorn, A. and
Ford, R.J.
Identification of a cDNA for a human high-molecular-weight B-cell
growth factor
Proc. Natl. Acad. Sci. U.S.A. 90 (13), 6330-6334 (1993)
JOURNAL MEDLINE 93317675
PUBMED 8327514
REFERENCE 2 (bases 1 to 1793)
Ambrus, J.L., Jr.
Direct Submission
Submitted (03-MAY-1993) Julian L. Ambrus, Jr., Department of
Medicine, Washington University School of Medicine, St. Louis, MO
63110, USA
COMMENT On May 25, 1995 this sequence version replaced gi:347805.
Original source text: Homo sapiens cDNA to mRNA.
FEATURES
Location/Qualifiers
1..1793
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
source

```

```

/cell_line="Namalva"
/tissue_type="Burkitt lymphoma"

```

ORIGIN

Alignment Scores:

```

Pred. No.: 6,18e-75 Length: 1793
Score: 1688.00 Matches: 339
Percent Similarity: 99.12% Conservative: 0
Best Local Similarity: 99.12% Mismatches: 3
Query Match: 62.47% Indels: 1
DB: 9 Gaps: 0

```

US-10-023-523-8 (1-530) x HUMHMBGCF (1-1793)

```

QY 189 AsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLys 208
Db 1791 AATTCACAGAGCAGATGAAGCTTCCTACAGAAAAGCAGAGCCAGCTGGTGCAAGAGAAG 1732

QY 209 AspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeu 228
Db 1731 GACCACCTGGCGGTGAGCAGCAGCAAGCC-GTCTGCGCCCGCAGCAAGCTTGAGAGCCTA 1673

QY 229 CysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGlyValGlnArgAlaArg 248
Db 1672 TCCGCTGAGCTGCACGGGCAACCGCTCCCTCAAGAGAAAGGTGTGACAGCGGCCCGG 1613

QY 249 GluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIle 268
Db 1612 GAGGAGGAGGAGAGCGCAAGAGGTGACCTCGACCTCCAGGTGACACTGAATGACATT 1553

QY 269 GlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGlu 288
Db 1552 CAGCTGCAGATGGAAACAGCACAATGAGCGCAACTCCAAAGTGGCCCAAGAGAACATGGAG 1493

QY 289 LeuAlaGluArgLeuLysLysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAsp 308
Db 1492 CTGCTCAGAGGCTCAAGAGCTGATTGACAGTATGAGCTGCGCGGAGGACATATCGAC 1433

QY 309 LysValPheLysHisLysAspLeuGlnGlnLeuValAspAlaLysLeuGlnAla 328
Db 1432 AAAGTCTTCAACACAAAGGACCTACAAACAGCAGCTGTGTGGATGCCAAGCTCCAGCAGGCC 1373

QY 329 GlnGluMetLysLysGluAlaGluArgHisGlnArgGluLysAspPheLeuLys 348
Db 1372 CAGAGATGTAAAGGAGGAGCAGAGAGCGGCACCCAGCGGAGAGAGGATTTCCTCTGAAA 1313

QY 349 GluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLeuLys 368
Db 1312 GAGGACGTAGAGTCCAGAGAGATGTGTGAGCTGATGAAGCAGCAAGAGAGACCCACCTGAAG 1253

QY 369 GlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSer 388
Db 1252 CAACAGCTTGCCCTATACACAGAGAAGTTTGAAGAGTTCCAGAACACACTTTCACAAAGC 1193

QY 389 SerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLys 408
Db 1192 AGCGAGGTATTCCACCATTCAGCAGGAGATGGAAAGATGACTTAAAGATCAAGAAG 1133

QY 409 LeuGluLysGluThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeu 428
Db 1132 CTGGAGAAAGAAACACCATGTACCGTCCCGTGGGAGAGCAGCAACAGGCCCTGCTT 1073

QY 429 GluMetAlaGluGluLysThrValArgAspLysGluLeuGluGluValLysLys 448
Db 1072 GAGATGGCTGAGGAGAAACACGTCCGGGATAAGAACTGGAGGGCTTGAGGTAAATATC 1013

QY 449 GlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArg 468
Db 1012 CAACGGCTGGAGAGCTGTCCGGGCACCTGCAGACAGAGCGCAATGACTGACACAGAGG 953

QY 469 ValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSerGlyProGluArgArg 488
Db 952 GTACAGGACCTGAGTGTGGTGGCCAGGGCTCCCTCACTGACAGTGGCCCTGAGAGGAGG 893

```

QY 489 ProGlyAlaProSerThrGluAlaProSerProArgValThrGluAlaProCysTyr 508
 Db 892 CCAGAGGGGCTGGGCTCAGACCCAGCTCCCGAGGGTCACAGAGGCTTGCTAC 833
 QY 509 ProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerIla 528
 Db 832 CCAGAGGACCCAGACAGAGCAGATCAGGCTGGGCTCAGAGCCACCTCCGCC 773
 QY 529 ArgAla 530
 Db 772 AGGCCC 767

RESULT 20
 AC108815
 LOCUS
 DEFINITION Mus musculus clone RP23-239013, WORKING DRAFT SEQUENCE, 5 unordered pieces.
 AC108815 GI:28882391
 VERSION HTG, HTGS, PHASE1, HTGS DRAFT.
 KEYWORDS Mus musculus (house mouse)
 SOURCE
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 181547)
 Birren, B., Nusbaum, C. and Lander, E.
 Mus musculus, clone RP23-239013
 2 (bases 1 to 181547)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Collangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Govette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Lakocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Melrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Trigilio, J., Vassiliev, H., Topham, K., Travers, M., Travis, N., Vo, A., Wyman, D., Ye, W.J., Young, G., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
 Submitted (31-JAN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 181547)
 Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C., MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Melrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,

Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (08-MAR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Mar 8, 2003 this sequence version replaced gi:28195423.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 Project Information
 Center project name: L20929
 Center clone name: 239 O.13

Summary Statistics
 Sequencing vector: Plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 180713 bases at least Q40
 Consensus quality: 180940 bases at least Q30
 Consensus quality: 181039 bases at least Q20
 Insert size: 172000; agarose-fp
 Insert size: 181147; sum-of-contigs
 Quality coverage: 8.0 in Q20 bases; agarose-fp
 Quality coverage: 7.6 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 15598: contig of 15598 bp in length
 15599: gap of 100 bp
 15699: contig of 1745 bp in length
 17443: gap of 100 bp
 17544: contig of 2265 bp in length
 39808: gap of 100 bp
 39809: contig of 36942 bp in length
 76850: gap of 100 bp
 76851: contig of 104597 bp in length.
 181547: contig of 104597 bp in length.

FEATURES

Location/Qualifiers
 1. 181547
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="RP23-239013"

misc_feature
 1. 15598
 /note="assembly_fragment"

misc_feature
 clone_end:SP6
 vector_side:left
 15699..17443

misc_feature
 /note="assembly_fragment"
 17544..39808

misc_feature
 /note="assembly_fragment"
 39809..76850

misc_feature
 /note="assembly_fragment"
 76951..181547

misc_feature
 /note="assembly_fragment"
 clone_end:T7
 vector_side:right

ORIGIN

Alignment Scores:
 Pred. No.: 3.38e-61 Length: 181547
 Score: 1452.00 Matches: 402

Percent Similarity:	42.59%	Conservative:	35
Best Local Similarity:	39.18%	Mismatches:	89
Query Match:	53.74%	Indels:	505
DB:	2	Gaps:	6
US-10-023-523-8 (1-530) x AC108815 (1-181547)			
QY	1	LysSerSerProGlyVclnProGluAlaGlyProGluGlyValaGlnGluAArgProSerGln	20
DB	110742	AGAGCAACCGGGGCAAGGGAACAGAGACTGGACGGAGCCATCAAAAGACCCAGACAG	110801
QY	21	AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProAArgLysProGlu	40
DB	110802	ACAGCTCTGGGGCAGAGCTGAAGT---TCCATAAGCCAGGTTCCTGGGAAGATGAG	110858
QY	41	GlyAlaGlnAlaAArgThrAlaGlnSerGlyAlaLeuAArgValSerGluGlyLeuSer	60
DB	110859	GGTCTCAGAGTAAAGCAGCGTAGACTGGGCTCTATGTGATGCTCTGAGGAGCTGAGC	110918
QY	61	ArgGlnLeuGluAspIleLeuSer---ThrTyCysValAspAsnAsnGlnGlyGlyProGln	80
DB	110919	CGSCAGTGGGAAGACATCTGATTAATAAATAATGTGTGAACAACATCGG-----GG	110969
QY	80	YcluaspglyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTy	100
DB	110970	AGACTCGGGCAGGA-GGTGAGCCCACTGAGCCGGAAGACATAGAGAAGTCTCTGAACCTA	111028
QY	100	rValAlaAArgAsnGlyGluProGluPro---ThrProValValTyGlyGluLysGluPr	119
DB	111029	TGCAGCCAGGAATGGGAAGCTGTAATCAGGCAATTCAGTAGTCAATGGCAAAAGGAGAC	111088
QY	119	oSerlyGlyAspProAsnThrGluGluAArgGlnSerAspGluValGlyAspArgAs	139
DB	111089	CTCTAAGGGGACGCTGTTCAGAAGAGATCCGAGCAAGTATGATCAGGTGGAGACTGGA	111148
QY	139	PhiSarArgProGln-----	144
DB	111149	CCATCGAAGGCAACAAGAAAGAAAGAACTCAAGAGGTGGACTTCAGAAAAATCAATAAC	111208
QY	144	-----	144
DB	111209	CCGTTAAAAAATGGGCTCAGAACTGAACAAAGGATTCACCTCAGGAAATACCGAATGG	111268
QY	144	-----	144
DB	111269	CAGAGAAGCACTGAAAAAATGTTCAACATCCTTAATCATCAGGGAATGCAAAATCAAAA	111328
QY	144	-----	144
DB	111329	CAACCTGAGATTCACCTCACACAGTCAAGATGGCTGAAGTCAAAAATTCAGGTGACA	111388
QY	144	-----	144
DB	111389	GCAGATGCTGGCGTGGATGTGGAGAAAGAGAAACACTCTCTCAATGTTGTGGGATTGCA	111448
QY	144	-----	144
DB	111449	GGCTTGTACACCACTCTGGAAATCAGTCTGGTGGTTCCTCAGAAAATTCGACACAGTAC	111508
QY	144	-----	144
DB	111509	TACTGGAGATCCCGCAATACCTCTCCTGGGCATATATCCAGAAGATGTCCCAACCGGTA	111568
QY	144	-----	144
DB	111569	AGAAGGACACATGCTCCACTATGTTTCATAGACGCTTATTTATAAGCCAGAAGTGGGA	111628
QY	144	-----	144
DB	111629	AAGAACCCAGATGCCCTCAACAGAGGAATGGATACAGAAAAATGTGTATCTCTACACA	111688
QY	144	-----	144

QY 218 aValLeuAlaAtrSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSe 238
 Db 112829 TGCTCTGGCCGAGCAACTTGAGAGCTGCTCTGGAGAGTGCACACACACACCGCATC 112888
 QY 238 rLeuLysGluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValTh 258
 Db 112889 CCTAATGGAAGAGTGTGTCAGGAGCCCATGAGAAGAGGAGGAGACCAAGAGAGTGAC 112948
 QY 258 rSerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluAr 278
 Db 112949 TTAACACTTCCAGTGACCTGATGACATTCAGTTGAGATGAGCAGCATTACTA-CC 113007
 QY 278 gAsnSerLysLeuArgGlnGlnLeuMetGluLeuAlaArgLeuLysLysLeuLeuGl 298
 Db 113008 AAAATCCAGCTGTGCTCAAGAGAAATATGAACTAGCCGAGAGGCTCAAGAAAGTTGATGA 113067
 QY 298 uGlnTyrGluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGl 318
 Db 113068 GCATACAGACTTCGTTAGAGCATATGCACA--GCTCTCAACATAAGGCAAGCAGCA 113125
 QY 318 nGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluGluAr 338
 Db 113126 GCAGCTAGTGGACACCAAGCTACTGTCAGGCCAGAAAGATGCTGAAGAGGCGAGAGGAGCA 113185
 QY 338 gHisGlnArgGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGl 358
 Db 113186 GCACCAAGAGGCGAAAGATTTCTCAGAGGAGAGCTGGAATCCAGAGAGATGCGCA 113245
 QY 358 uLeuMetLysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeu--TyrThrGluLys 377
 Db 113246 GCTGATAAAACAGCAGAGACCACTTAAGAGCAGCTTTGCTGCGGCTACACGAGAGAG 113305
 QY 378 Phe-GluGluPheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGl 397
 Db 113306 TTCGAGGAGTTCACAGACACACTTTCAGAGAGCTGAAGTGTACACGTTCAACA 113365
 QY 397 nGluMetGluLysMetThrLysLysLysLysLysLysGluLysGluThrThrMetTyrAr 417
 Db 113366 CGAGATGGAAGAGATGACAAAGAAATCAAGAGAGTGGAGAAAGAAACCAACCATGTACCG 113425
 QY 417 gSerArgTrpGluSerSerAsnLysAlaLeu-LeuGluMetAlaGluLysThrVala 437
 Db 113426 ATTCCGTTAGGAAGCAGACACAGCTCTGCTGGAGATGGCTGAGAGAAACCCCTCC 113485
 QY 437 rAspLysGluLeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuCysArgA 457
 Db 113486 TGGACAAATTTGATGGAAGGCTCGCGGTGAAATCCG--CGGTTGGAGAGCTGTGCGGAG 113544
 QY 457 lAlaGlnThrGluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyG 477
 Db 113545 CGCTGAGACTGAGTCAATGACCTTCAAGAGGATATAGGACATGCTGCTGGGGCT 113604
 QY 477 lGlySerLeuThrAsp-----SerGlyProGluArgArg 488
 Db 113605 AGGGCTCCCTCACTGCTGTCAGAGCTGCGAGGCTGCCCTGCTCAAGGACAGGT 113664
 QY 489 ProGluGlyProGlyValAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyr 508
 Db 113665 GTTGAGTCTCTGGGGCTCAGCCATCCATCTCTCAAGGGGCAACAGAGCTCCCGACTGC 113724
 QY 509 ProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAla 528
 Db 113725 CCAGGAACATCGATCACAGAAACAGACGACCAACCG--CCCGAGAGGCCCTCTGCTGCC 113782
 QY 529 ArgAla 530
 Db 113783 ATTGCC 113788
 RESULT 21
 AX878120
 LOCUS
 DEFINITION Sequence 13025 from Patent EP1074617.

ACCESSION AX878120
 VERSION AX878120.1 GI:40032856
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 REFERENCE 1
 AUTHORS Ota, T., Isega, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
 Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
 TITLE Primers for synthesizing full-length cDNA and their use
 JOURNAL Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Research Association for Biotechnology (JP)
 LOCATION/Qualifiers
 1. 2045
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 52. 1638
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAE90020.1"
 /db_xref="GI:40032856"
 /translation="MATRVEAARGGCGGAEATEAGRGRRSPROKTEIGTMEBAG
 1CGLVADMLCNSQNDILHOGSCNGGTNNKSHLEEDGSDFTITENLUSPAYCT
 QESREELFGGEARTDPDGGQDSECRNKEKTIGKSVLLMOALNTLSPEKLAALC
 KYADLLEERSVOKMKILCKQAQIVKEKHQSHSKAILARLSKLSLCELORH
 NKTLEENMQOAREEERREKETAHFQITLDEIOAQLEQHDHNAKLRENIELEKEL
 KKLLEVALREEDIDVFKEKEIQQLQDLAQTTQLTKEADEGHEQREBELKEAT
 ESRHKYQMKQOEVQLKQQLSLYMDKFEFEFTMAKSNELFTTFREMEKMTKKL
 EKETIIWRTWNNKALLQWABEKTVRDEYKALQIKLERLEKLCRALQTERNELNE
 KVEVLKQVSTKAAIKANRDLATPVQPCALDLSKELNTSSKALGAHLBAEPKQ
 RSAVQKPESTGSAPIAESVD"
 ORIGIN
 Alignment Scores:
 Pred. No.: 5,17e-55 Length: 2045
 Score: 1285.00 Matches: 284
 Percent Similarity: 67.49% Conservative: 71
 Best Local Similarity: 53.99% Mismatches: 132
 Query Match: 47.56% Indels: 39
 DB: 6 Gaps: 8
 US-10-023-523-8 (1-530) x AX878120 (1-2045)
 QY 12 GluGlyAlaGlnGluArgProSerGlnAlaProAlaValGluAlaGluGlyProGly 31
 Db 70 GAGGCGAGCGCGGAGAGAGCGCGCGCCGCGGAGCGGAGCGGCGGCGGCGGCGG 129
 QY 32 SerSerGlnAlaProArgLys-----ProGluGlyAlaGlnAlaArgThrAlaGlnSer 49
 Db 130 CGGCGAGCGAGCGCGCGGAGAGAGTTTGAATTTGGCAATGGAAGAGCGTGAATTTGT 189
 QY 50 GlyAlaLeuArgPvalSerGluLeuSerArgGlnLeuGluAspIleLeuSerThr 69
 Db 190 GGGCTAGGGGTGAAGCAGATATGTTCTGTAACTCTCAATCAATCATATTCTTCAACAT 249
 QY 70 TyrCysValaspAsnAsnGlnGlyGlyProGlyGluAspGlyAlaGlnGlyProAla 89
 Db 250 CAA-----GGCTCAAAATTTGGTGGCAGCAAGTAAGCAAGCAATTTCGAGAGGATGAA 303
 QY 90 GluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsnGlyGluProGluPro 109
 Db 304 GGCAGTGACTTATAACAGAGAACAGGAAATTTGGTGAGC-----CCA 345
 QY 110 ThrProValValTyrGlyGluLysGluProSerLysGlyAspProAsnThrGluGluLle 129
 Db 346 GCATACCTGCAGCGCAAGAAATCAAGAGAGGAAATCCCTGGGGGA-----GAAGCT 393
 QY 130 ArgGlnSerAspGluValGlyAspArgAspHisArgArgProGlnGluLysLysAla 149
 Db 394 CGAAGACATCCCTGATGCTGAGAGATTCAGAGTGCAGACAGGAGCAAGAGAGAAACT 453

QY 150 LysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrPro 169
Db 454 -----TTAGGAAGAAGATTATTATCTGATGCAAGCCCTTAACACCCCTTCAACCCCA 507
QY 170 GluGluLysLeuAlaLeuCysLysLysTyrAlaGluLeuGluGluHisArgAsn 189
Db 508 GAGGAGAAGCTGGCAGCTCTCTGTAGAAATATGCTGATCTTCTGGAGCAGCAGGAGT 567
QY 190 SerGlnLysGlnMetLysLeuGlnLysLysGlnSerGlnLeuValGlnGlnLysAsp 209
Db 568 GTTCAGAGCAATGAAGATCTCGAGAGAGCAAGCCAGATGTGTAAGAGAAAGATT 627
QY 210 HisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLysCys 229
Db 628 CACTTGCAGAGTGAACATAGCAGCTATCTTGGCAAGCAAGCTAGATCTCTTTC 687
QY 230 ArgLeuLeuGlnArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGlu 249
Db 688 AGAGAACTTACGCTGCAATAAGACGTTAAGAGAGAAATATGCGAGCAGCAGAG 747
QY 250 GluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGln 269
Db 748 GAAGAAGAACGCTTAAGAGCAACTGCATTTCCAGATACCTTAGATGAATTCAA 807
QY 270 LeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGlnAsnMetGluLeu 289
Db 808 GCCAGCTGGAGCAGCATGCATCCACAAAGCCAACTCCACAGCAAAACATTGAGCTG 867
QY 290 AlaGluArgLeuLysLysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAspLys 309
Db 868 GGGGAGAAGCTTAAGAGCTCTATGACAGTACCCTGAGGAGAGCAATTTGATAAG 927
QY 310 ValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGln 329
Db 928 GTGTTCAAACTGAAGAACTGCAACAGCAGCTCTGTCGAACTGCAAGCAACGACA 987
QY 330 GluMetLeuLysGluAlaGluGluArgHisGlnArgGluLysAspPheLeuLysGlu 349
Db 988 CACTGTATAAAGAGCTGATGAATAACATCAGAGAGAGAGTGTATTTATTAAGA 1047
QY 350 AlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLeuLysGln 369
Db 1048 GCGACAGATCGAGGCAAAATACGAAATAATGAACAGCAGAGAGTACAACTAAGACAG 1107
QY 370 GlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSer 389
Db 1108 CAGCTTCTCTTATATAGGATAGTTTGAAGATTCAGACTACCATGGCAAAAGCAAT 1167
QY 390 GluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysIleLysLysLeu 409
Db 1168 GAAGTGTTCACCTTCACAGAGAAATGGAAGATGCAAGAAATTAAGAACTG 1227
QY 410 GluLysGluThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeuGlu 429
Db 1228 GAAAGAAACAATAATTTGGCTTACCAATGGGAAACAAATAAAGCACTTCTGCAA 1287
QY 430 MetAlaGluGluLysThrValArgAspLysGluLeuGluGluValLysIleGln 449
Db 1288 ATGGCTGAAGAGAAACAGTCCGTGTATAAGATGACAGGCCCTTCAATAAAGCTG 1347
QY 450 ArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgVal 469
Db 1348 CGGTTAGAGAAGCTGTGAGGGCTCTTCAACAGAAAGGAATGAGCTCAATGAGAGGTG 1407
QY 470 Gln-----AspLeu 472
Db 1408 GAAGTCTGTAAGAGCAGGATCCATCAAGCGGCCATCAAGCGCGCAACAGGATTTA 1467
QY 473 SerAlaGlyGlnGlySerLeuThr-----AspSerGlyProGluArgArgProGlu 490
Db 1468 GCAACACCTGTGATGAGCCCTGCTACTGCCTGATTTCTCAACAGAGCTGACACTTCC 1527
QY 491 GlyPro-----GlyAlaGlnAlaProSer-SerProArgValThrGluAlaProCys 507

Db 1528 TCAGAAAGAGCCCTGGAGCGCACCTGGAGCTGAGCCCAAGAGTCCAGAGAGCGCTGTG 1587
QY 507 sTyrProGlyAlaPro 512
Db 1588 CAAAAGCCCCCGTCCA 1603
RESULT 22
BD157034
LOCUS BD157034 2045 bp DNA linear PAT 17-JAN-2003
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD157034
VERSION BD157034.1 GI-27862792
KEYWORDS JP 2002191363-A/11877.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2045)
AUTHORS Ota.T., Isogai.T., Nishikawa.T., Hayashi.K., Saito.K., Yamamoto.J.,
Teshi.S., Sugiyama.T., Wakamatsu.A., Nagai.K. and Otsuki.T.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002191363-A 11877 09-JUL-2002;
HELIX RESEARCH INSTITUTE
COMMENT OS Homo sapiens (human)
PN JP 2002191363-A/11877
PD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
PI SAITO,
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI KEIICHI NAGAI, TETSUJI OTSUKI
PC
C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
10,
C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers (52)..(1635).
FT CDS
FEATURES
source
1..2045
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 5,17e-55 Length: 2045
Score: 1285.00 Matches: 284
Percent Similarity: 67.49% Conservative: 71
Best Local Similarity: 53.99% Mismatches: 132
Query Match: 47.56% Indels: 39
DB: 6 Gaps: 8
US-10-023-523-8 (1-530) x BD157034 (1-2045)
QY 12 GluGlyAlaGlnGluArgProSerGlnAlaAlaProAlaValGluAlaGluGlyProGly 31
Db 70 GAGGAGCGCGGGGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 129
QY 32 SerSerGlnAlaProArgLys-----ProGluGlyAlaGlnAlaArgThrAlaGlnSer 49
Db 130 CGCGACCGCAGCG 189
QY 50 GlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGluAspIleLeuSerThr 69
Db 190 GGGCTAGGGGTGAAGCAGAGATATGTTGTAACTCTCAATCAATGATATTTCTTCAACAT 249
QY 70 TyrCysValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAlaGlnGlyGluProAla 89
Db 250 CAA-----GGCTCAAAATTTGGTGGCACAAGTAAACAGCATTCATTGGAAGAGATGAA 303
QY 90 GluProGluAspAlaGluLysSerArgThrTyrValAlaArgAsnGlyGluProGluPro 109

```

Db 304 GGCAGTGACTTATACAGAGACAGGAATTTGGTGAGC-----CCA 345
QY 110 ThrProValValTyrGlyGluLysProSerLysGlyAspProAsnThrGluGluLeu 129
Db 346 GCATACCTGCAGCAGCAAGATCAAGAGAGGAATCCCTGGGGGA-----GAAGCT 393
QY 130 ArgGlnSerAspValGlyValAspArgAspHisArgArgProGlnGluLysLysVala 149
Db 394 CGAACAGATCCCTCGATGATGAGCAGGATTCAGAGTGCACAGCAAGCAAGAAAGAACT 453
QY 150 LysGlyLeuGlyLysGluLeuThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrPro 169
Db 454 -----TTAGGAAAAGAAAGTTTATTACTGATGCAAGCCCTAAACACACCTTTTCAACCCCA 507
QY 170 GluGluLysLeuAlaLeuLysLysLysLysLysLysLysLysLysLysLysLysLys 189
Db 508 GAGGAGAGCTGGCAGCTCTCTGTAGAAATATGCTGATCTCTGAGGAGAGCAGGAGT 567
QY 190 SerGlnLysGlnMetLysLeuLysGlnLysLysGlnSerGlnLeuValGlnGluLysAsp 209
Db 568 GTTCAGAGCAATGAAGATCTCTGAGAGAGCAAGCAAGCCAGATTTGTAAGAGAGAAATG 627
QY 210 HisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCys 229
Db 628 CATTTCAGAGTGAACATAGCAGGCTATCTTGGCAAGAGCAAGCTAGAAATCTTTTTC 687
QY 230 ArgGluLeuGlnArgHisAsnArgSerLeuLysGluGlyValGlnArgAlaArgGlu 249
Db 688 AGAGAACTTCAGCGTCACATAAGAGCTTAAAGGAGGAAATATATGACAGCAGCAGAG 747
QY 250 GluGluLysArgGlyGluValThrSerHisPheGlnValThrLeuAsnAspIleGln 269
Db 748 GAAGAGAGCAGCGTAAAGAGCACTGCACATTTCCAGATACCTTTAGATGAATTCAA 807
QY 270 LeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluLysMetGluLeu 289
Db 808 GCCAGCTGAGCAGCATGACATCCACACGCAACTCCGACAGGAAATCATTTAGCTG 867
QY 290 AlaGluArgLeuLysLysLeuGluGlnTyrGluLeuArgGluGluHisLysAspLys 309
Db 868 GGGAGAGCTAAAGAGCTATCTGAAACAGTACGCACTGAGGAGAGCAGCATTTGATAAG 927
QY 310 ValPheLysHisLysAspLeuGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGln 329
Db 928 GTGTTCAACCGTAAGCACTGCAACAGCAGCTCGTGGATGCCAACTGCAGCAACAGCA 987
QY 330 GluMetLysGluAlaGluArgHisGlnArgGluLysAspPheLeuLysGlu 349
Db 988 CMACTGATAAAGAGCTGATGAACACATCAGAGAGAGAGAGTATTTTATTAAAGNA 1047
QY 350 AlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnThrHisLeuLysGln 369
Db 1048 GCGACAGATTCAGGCACAAATACGAAACAAATGAACAGCAGGAGTACAACTAAACAG 1107
QY 370 GlnLeuAlaLeuTyrThrLysLysPheGluGluPheGlnAsnThrLeuSerLysSer 389
Db 1108 CAGCTTCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1167
QY 390 GluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLeu 409
Db 1168 GAACTGTTTACAACTTCAGACAGCAAGTAAAGAGTGAACAGAAATGAACAGAAATTA 1227
QY 410 GluLysGluThrThrMetTyrArgSerArgTTPGluSerSerAsnLysAlaLeuGlu 429
Db 1228 GAAAGAGAAACAAATATTTGGGCTACCAATGGGAAACAAATATAAGACCTTCTGCA 1287
QY 430 MetAlaGluGluLysThrValArgAspLysGluLeuGluGlyLeuGlnValLysIle 449
Db 1288 ATGGCTGAAGAGAAACACAGCTCGCTGATAAAGAGTCAAGAGCCCTTCAATAAAG 1347
QY 450 ArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArg 469

```

```

Db 1348 CGTTTAGAAGCTGTGCGAGGCTCTTCAACAGAAAGAAATGAGCTCAATGAGAGGTG 1407
QY 470 Gln-----AspLeu 472
Db 1408 GAAGTCTGAAAGACAGGTATCCATCAACAGCGCCATCAAGCGGCAACAGGATTTA 1467
QY 473 SerAlaGlyGlyGlnGlySerLeuThr-----AspSerGlyProGluArgArgProGlu 490
Db 1468 GCAACACCTGTGATGCGAGCCCTGTACTCCCTGCTGATTTCTCAGAGGAGCTCAAC 1527
QY 491 GlyPro-----GlyAlaGlnAlaProSer-SerProArgValThrGluAlaProCys 507
Db 1528 TCAGAAAGAGCCCTGGGAGCGCACCTGAGGCTGAGCCCAAGAGTCAGAGAGCGGTG 1587
QY 507 STyrProGlyAlaPro 512
Db 1588 CAAAGCCCCCGTCCA 1603

RESULT 23
AK002071 2045 bp mRNA linear PRI 01-AUG-2002
LOCUS Homo sapiens cDNA FLJ11209 fis, clone PLACE1007946.
DEFINITION AK002071
ACCESSION AK002071
VERSION AK002071.1 GI:7023730
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y., Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H., Wagatsuma, M., Hosoi, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takamashi, M., Chiba, Y., Ishida, S., Murakawa, K., Ono, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahara, K., Masuho, Y., Ninomiya, K. and Iwayanagi, T.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2045)
AUTHORS Isogai, T. and Otsuki, T.
TITLE Direct Submission
JOURNAL Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 252-0812, Japan (E-mail: genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.
FEATURES
source
location/Qualifiers
1..2045
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE1007946"
/tissue_type="placenta"
/clone_lib="PLAC1"
/notes="cloning vector: pME18SFL3"
52..1638
/notes="unnamed protein product"
/codon_start=1
/protein_id="BA092068.1"
/db_xref="GI:7023731"
/translation="MATRVEAARGCGGAEATEARGRRRSRPRKTEIGTWERAG
ICGVKADMLCNSQNDILQHQSCNGGTSNKHSELEDEGSDFTFENRNLSPAYCT
QESSEIIPGEARTDPDQDSECNRNKEKTLGKREVLMLQALNTLSTPEKLAALC
KYADLLEERSVQKQMKILQKQAGIVKEKHLQSHKXAILARSLKLSLELQKH
NKTLKENMQOAREEERREKTAHFQITLDEIQALQLEQIHNAKLQENIBELGKL
KKLLEQVALREEHIDKVKFKELQQLQVDAKLOQTTLIKEADEKHQREERFLKEAT
ESRHKYQMKQEQVLAQQLSLYMDKFEFQTTMAKSNELFTTFRCEMEQWTKKKL"
CDS

```


(DKFZp451B226) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further information about the clone and the sequencing project is available at <http://mips.gsf.de/proj/cDNA/>.

FEATURES

source
location/Qualifiers
1..4625
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp451B226"
/tissue_type="human skeletal muscle"
/clone_lib="451 (synonym: hicc1). Vector pSport1; host DH10B; sites NotI + SalI"
/dev_stage="adult"
4583..4588
polya_signal
polya_site
4605

ORIGIN

Alignment Scores:

Pred. No.: 9.95e-54 Length: 4625
Score: 1266.50 Matches: 276
Percent Similarity: 61.51% Conservative: 90
Best Local Similarity: 46.39% Mismatches: 134
Query Match: 46.87% Indels: 95
DB: 9 Gaps: 10

US-10-023-523-8 (1-530) x HSM807779 (1-4625)

QY 18 ProSerGlnAlaAlaProAlaValGluAlaGlu----- 28
Db 71 CCTTCTGATCTACTCCAGAGATGGAGGCTTAATCACTCTGAACAGCTCTCAGCGGAACGA 130
QY 29 -----GlyProGly-----SerSerGlnAlaPro-----ArgLysPro 39
Db 131 CAGTCAACACCTCCAGGTGACAGTTCATCATTAACCCAGTCACAAATGCCCTGGAGAGGAA 190
QY 40 GluGluAlaGlnAlaArgThr-----AlaGlnSerGlyAlaLeuArg 53
Db 191 GATGGCCAGAGATTCTCAACCCCGTCCACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 250
QY 54 AspValSerGluGluLeuSerArgGlnLeuGluAlaSpileLeuSerThrTyrCysValAsp 73
Db 251 GATATCTCTGAAGAGCTGAATCGACAGCTGGAAGACATCATTAACATTAT-----GGG 304
QY 74 AsnAsnGlnGlyGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluProGluAsp 93
Db 305 TCTGCTGCCAGCACAGCAGGGAAGAGGGCTCTGCCAGGGCCAGTCAGCAGCCTGAGAA 364
QY 94 AlaGluLysSerArgThrTyrValAlaArgAsnGlyGluProGluProThrProValVal 113
Db 365 GCAGATACCTGACAC-----GAGGATGGGACTGTGAGGAACACTCAAGAG 415
QY 114 TyrGlyGluLysGluProSerLysGlyAspProAsnThrGluGluLeuArgGlnSerAsp 133
Db 416 GCTGGAAGAGAACCCCTTCTCTGAGAGCCACCCTCTGTCAAA-----GAG 463
QY 134 GluValGlyAspArgAspHisArgArgProGlnGluLysLysAlaLysGlyLeuGly 153
Db 464 CCGCTCAGCAATAAGAGAGCAAAA-----TTGGAAGAAATAATCTTAAGGATTAGGC 517
QY 154 LysGluLeuThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeu 173
Db 518 AAAGAGCCCAACCTGTAATGCAAAATCTGAACAGATTGCAAAACACCGGAGAAAGTTT 577
QY 174 AlaAlaLeuCysLysLysTyrAlaGluLeuLeuGluHisArgAsnSerGlnLysGln 193
Db 578 GATTTTATTCAAGAAGATATCTGAATTCGTGGATGAACATCGTACTGAGCAAAAGAAAG 637
QY 194 MetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGly 213
Db 638 TTAAAGCTCTCCAAAGAAACAGGTACAAATTCAAAAGAAAGAACAGACCAAGTTACAGGT 697

QY 214 GluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGln 233
Db 698 GAACACAGCAGAGCTATCTCTGCTGAAGCAAAATGGAGAGTCTGTGCGGAGAGCTGAG 757
QY 234 ArgHisAsnArgSerLeuLysGluGlyValGlnArgAlaArgGluGluGluLys 253
Db 758 AGACACAACAAGACTCTGAAGAAAGAGCGCTTCAGCGGGCAGCTGAGGAAGAGAGAA 817
QY 254 ArgLysGluValThrSerHisPheGlnValThrLeuAsnAspLleGlnLeuMetGlu 273
Db 818 AGGAAGGAATACCAAGCCATTCAGAGTACCTTCAGGACATCCAGGGCCAGATCGAG 877
QY 274 GlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeu 293
Db 878 CAGCAGAGTGAAGCAATATGAGCTCTGTCAGGAGAACACAGAGCTTGCAGAAAAGCTG 937
QY 294 LysLysLeuLleGluGlnTyrGluLeuArgGluHisLleAspLysValPheLysHis 313
Db 938 AAAAGCATCATGATCATGATGATGAGCTCAGAGAGAGCATCTGGACAAAATATTTAAACAC 997
QY 314 LysAspLeuGlnGlnLeuValAlaAlaLysLeuGlnGlnAlaGlnGluMetLeuLys 333
Db 998 AGAGNACTGTCAGCAGAGAGTCTGTCGATGCAAAAGCTTGAAGAGCCCAAGAAATATGAG 1057
QY 334 GluAlaGluGluArgHisGlnArgGluLysAspPheLeuLysGluAlaValGluSer 353
Db 1058 GAAGCGAGAGCGACACAAACGAGAAAAAATATTTGCTGAACACGAGCAGAGAGTGG 1117
QY 354 GlnArgMetCysGluLeuMetLysGlnGlnThrHisLysLysGlnGlnLeuAlaLeu 373
Db 1118 AAATCTGAGCGAAATGCTGAAGAGCAAGAGACAGTCTCTGACAGGCTCAGCTCCTCTC 1177
QY 374 TyrThrLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluValPheThr 393
Db 1178 TACTCAGNAGGTTTGAAGAATTCAGAGCAGCACATACTAAAGACACGAGGTGTTTGC 1237
QY 394 ThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLysLysLysLysLys 413
Db 1238 ACGTTCAACAGGAAATGGCAAAACAACTAAGAAATGAAGAGCTGGAAGAGAGACACA 1297
QY 414 ThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeuGluMetAlaGluGlu 433
Db 1298 GCCACATGGAAGCCCATTTGAGAACTGTAAACAAGCTCTGTGGACATGATTGAGAG 1357
QY 434 LysThrValArgAspLysGluLeuGluGlyLeuGlnValLysLysLysLysLysLys 453
Db 1358 AAACACTGAGAGTAAAGAATATGATGCTTTGTGATGAATAATCGGAGGCTAGAGAAC 1417
QY 454 LeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgValGlnAspLeuSer 473
Db 1418 CTCTGCGCTCTTTACAGAGAGAGAGAACTCAAAAATAATCAGAGAGAGAGAGAG 1477
QY 474 AlaGlyGlnGlySerLeuThrAspSerGlyProGluArgArgProGluGly----- 491
Db 1478 ATATCTGAAGAGATGACCAAGTCCAGCACAACCTCCGATGAGAGCCAGAGTCAAACGTC 1537
QY 491 ----- 491
Db 1538 TCTGTGATCAAGAGATTGACGAGAGAGGTTAATAGTGTCAAAACCCGCTGAAAAAT 1597
QY 491 ----- 491
Db 1598 CTGCGCACACCTTCATGATAATTATCATCATCAGAGTCAACCCCGCAGCTCCAAAGAA 1657
QY 492 -----ProGlyAlaGlnAlaProSer 498
Db 1658 ACCCAACCCGAAATAGGAGATCTCAGAGAGAGTGTGACCGCGCTCTCAGAGAGAGAG 1717
QY 499 SerPro-----ArgValThrGluAlaProCysTyrProGlyAlaProSer 513
Db 1718 CAACCCCTCTGATCCCTTTCAGCGGATTCAGAGAGTCCCTGCTCCCTCCCTTCTCTAG 1777
QY 514 ThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAla 528

538 CCGCTCAGCAATAAGCAGCAAAA-----TTGGAAGAGAAAATCTTAAAGGATTAGGC 591
154 LysGluIleThrLeuLeuMetGlnThrLeuSerThrProGluGluLysLeu 173
592 AAAGAGCAAACTGCTAATGCAAAATCTGAACACAGTTTGCACCAACCGGAGAAAGTTT 651
174 AlaAlaLeuCysIleLysValAlaGluLeuGluGluHisArgAsnSerGlnLysGln 193
652 GATTTTATTCAGAAAGTATCTGTAATTCGTGATGAACATCTACTCAGCAGAAAGAG 711
194 MetLysLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGly 213
712 TTAAGCTCTCTCAAAAGAAACAGGTACAAATTCAAAAGAAAGAACGACGATTACAAGT 771
214 GluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGln 233
772 GAACACAGCAGACTATCTCGCTCGAAGCAAAATTCGAGAGCTGTGCGCGAGCTCGAG 831
234 ArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGluGluLys 253
832 AGACACACAGACTCTGAAGGAGAGGCGCTTCAGCGGCACGTCGAGGAAGAGAA 891
254 ArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGlu 273
892 AGAAGCAATCAACAGCAATTTCCAGAGTACCTCAGGACATCCAGGCGCAGATCGAG 951
274 GlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeu 293
952 CACAGAGTGAGCAATATCAAGCTCTGTACAGAGACACAGAGCTTCGAGAAAGCTG 1011
294 LysLysLeuIleGluGlnValGluLeuArgGluGluHisLysLysValPheLysHis 313
1012 AAAAGCATCATCATGATGATGCTCAGAGAGGAGCATCTGCAAAATATTAAACAC 1071
314 LysAspLeuGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMetLys 333
1072 AGAAGCTCAGCAGAGAGCTGTGGATGCAAGCTTCAGAGAGCCCAAGAAATCATGAG 1131
334 GluAlaGluGluArgHisGlnArgGluLysAspPheLeuLeuLysGluAlaValGluSer 353
1132 GAAGCGAGGAGGACACAAACAGAGAAAGAAATATTGCTGAACAGCAGCAGCAGAGTGG 1191
354 GlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLysLysGlnGlnLeuAlaLeu 373
1192 AAATTCAGGCGAAGTGTCAAGGAGCAAGAGACATCTCGAGGCTCAGCTCATCTC 1251
374 TyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluValPheThr 393
1252 TACTCAGAGCTTTGAAGAAATTCACAGCAGACACTAACTAAAGCAACGAGGTGTTGCC 1311
394 ThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLysLysLysLysLys 413
1312 ACCTTCAACAGGAAATCGACAAACAACTAAGAAATGAAGAGCTGGAAGAGGACACA 1371
414 ThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeuMetAlaGluGlu 433
1372 GCCACATGGAAGCCCCATTTGAGAACTGTACAAAGCTCTGTGGACATGATGAAGAG 1431
434 LysThrValArgAspLysGluLeuGluGluValLysLysLysLysLysLysLysLys 453
1432 AAAGCACTGAGAGCTAAAGAAATATGAGTGTCTTGTGATGAAATCGGAGGCTAGAGAC 1491
454 LeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgValGlnAspLeuSer 473
1492 CTCTGCCGTGCTTTACAGAGAGAGAGAAACCAATCCCAAAAAAATACAGAGCGCAGAA 1551
474 AlaGlyGlyGlnGlySerLeuThrAspSerGlyProGluArgArgProGluGly----- 491
1552 ATATCTGAAAGGATGACCAAGTCAGCACAACCTCCGATGAGAGCCAGAGTCAACAGTC 1611
491 ----- 491
1612 TCTGTGGATCAAGAGATTGACGAGGAGGTTAATAGTGTCCAAACCGCGGTGAAAAAT 1671

491 ----- 491
1672 CTGCCACAGCTTCATGATAATTTCATCATCCAGAGTCAACCCGCGACCAAGTCCAAAGAA 1731
492 -----ProGlyAlaGlnAlaProSer 498
1732 ACCCAACCCGAAATAGGAGTCTCTCAGGAGAGTCTGACCGCGCTCTCAAGGAGCCAGAG 1791
499 SerPro-----ArgValThrGluAlaProCysTyrProGlyAlaProSer 513
1792 CAACCCCTCTGATCCCTTCACGGGATTCAGAGAGTCCCTCGCTCCCTCAACTCCTCAG 1851
514 ThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAla 528
1852 GCTGAAGCCGAGGAGGAGGAGTGTGTAACCTCCCTCCCAAGGCC 1896
RESULT 27
HSM805276 4631 bp mRNA linear PRI 17-JUN-2003
LOCUS Homo sapiens mRNA; cDNA DKFZp451F022 (from clone DKFZp451F022);
DEFINITION complete cds.
ACCESSION AL834248
VERSION AL834248.2 GI:30268572
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4631)
AUTHORS Wambutt, R., Heubner, D., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
Fobo, G., Han, M. and Wiemann, S.
DIRECT SUBMISSION
SUBMITTED (17-JUN-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
COMMENT On Apr 30, 2003 this sequence version replaced gi:21739792.
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
sequenced by AGOVA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.
This clone (DKFZp451F022) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.

FEATURES
source

1..4631
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp451F022"
/db_xref="taxon:9606"
/map="6q24.1"
/clone="DKFZp451F022"
/tissue_type="human skeletal muscle"
/clone_lib="451 (synonym: hiccl). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
1..4631
/gene="DKFZp451F022"
1..2148
/gene="DKFZp451F022"
/note="similar to muscle-derived protein MDP77 variant
1 (Mus musculus), differentially spliced"
/cdon_start=1
/product="hypothetical protein"
/protein_id="CAD3924_2"
/db_xref="GI:30268573"
/translation="GCCSDLLFPASFLSLFLSPDTPPKMEANHSQLSAER
QSTPGDSSLSPLSHNGLEKEDQDPTVPPEKEASVHPDISELNQLEIDINTY
SAATAGKESGARASEQENASPDNEKDETTETAEAGREVPASGPTVKPEVSNK
EOKLEKILGLGKLANLLMONKLOTPEKFDLFKVAELLDEHTEQKKLLQ
KKVQIQEKDQLQSEHSRAILLARSLKSLCLRELQRHNKLTKEALQRAEEERKE

gene

CDS

ITSHFQSLTDIOGIEBQSERNNKLCQENTELAELKLSIDQYELREELHDXIFKRR
 ELQKLVDALEQAEKMEAEERHREKKEVLLNQAEMKLOAKVLKEQETVLOAULT
 LYSGRFEFQSTLTQKSNVETATPKQEMDKTKKMKLEKDTATWKARFENCNKALLDM
 IEEKALRAKEVECFVMIKIGLENLCRALQERNEHLKJEDAEISKDDOSOHNSDEE
 PESNVSDQDEIDAEEVNSVOTAVNKLATAFWIHHPESTPHQSKETQPEIGSSQESAD
 AALKEPQPLIPRDSSESLPPLTPQAEAGGSDAEPPSKASNSPAGLGAETQCSGL
 PVGAQAOASWKEPABEASGAPQAPTEASLQKMEADVAPACAAEESHVAAMPACEPS
 RQPPRAAEELPVGASAGPQPRNVADTNLEGVN
 4589.4594
 /gene="DKZp451F022"
 4612
 /gene="DKZp451F022"

polya_signal

polya_site

ORIGIN

Alignment Scores:

Pred. No.: 1.12e-53 Length: 4631
 Score: 1265.50 Matches: 276
 Percent Similarity: 61.51% Conservative: 90
 Best Local Similarity: 46.39% Mismatches: 134
 Query Match: 46.84% Indels: 95
 DB: 9 Gaps: 10

US-10-023-523-8 (1-530) x HSM805276 (1-4631)

18 ProSerGlnAlaAlaProAlaValGluAlaGlu-----28
 73 CTTCTGATCTACTCCCAAGATGAGGCGTAATCACTCTGAACAGCTCTCAGCGGAACGA 132
 29 -----GlyProGly-----SerSerGlnAlaPro-----ArgLysPro 39
 133 CAGTCAACACTCCAGGTGACAGTTTCATCTTACCAGTCACAAATGGCTGGAGAGGAA 192
 40 GluGlyAlaGlnAlaArgThr-----AlaGlnSerGlyAlaLeuArg 53
 193 GATGGCCAGGATTTCTCAACCCAGTCCCAACCCAGAGAAAGAGCAAGTGTGCAACCC 252
 54 AspValSerGluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysValAsp 73
 253 GATATCTCTGAGAGCTGATCGACAGCTGGAAGACATCATTAACACTTAT-----GGG 306
 74 AsnAsnGlnGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluProGluAsp 93
 307 TCTGCTGCCAGCACAGCAGGGAAGAGGGCTCTGCCAGGCGCCAGTCAGCAGCCTGAGAAT 366
 94 AlaGluLysSerArgThrTyrValAlaArgAsnGlyGluProGluProThrProValVal 113
 367 GCAGATCACTCACAAC-----GAGGATGGGACTGTGGAGAAACAACCTGAAGAG 417
 114 TyrGlyGluLysGluProSerLysGlyAspProAsnThrGluGluIleArgGlnSerAsp 133
 418 GCTGGAAGAACCCGTTGCTTCTGAGAGCCACCCTCTGTCAA-----GAG 465
 134 GluValGlyAspArgAspHisArgAArgProGlnGluLysLysAlaLysGlyLeuGly 153
 466 CCGCTCAGCAATAGGAGCAAAAA-----TTGAAAAAGAAATCTCAAAAGGATTAGGC 519
 154 LysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeu 173
 520 AAAGAGCCCAACCTGCTAATGCAAAATCTGACAACTTCGCAACACCCGGAAGAAAGTTT 579
 174 AlaAlaLeuCysLysLysTyrAlaGluLeuLeuGluGluHisArgAsnSerGlnLysGln 193
 580 GATTTTATTCAGAAAGTATGCTGAATTCGTGAATTCAGTGAACCTGACTGAGCAAAAGAG 639
 194 MetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGlnLysAspHisLeuArgGly 213
 640 TTAAGCTCTCTCAAAAGAAACAGGTACAAATTCAAAAAGAAAGAACGACCTTACAGGT 699
 214 GluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGlnLeuGln 233
 700 GAACACAGACAGACTATCTCGTCGTCGAAGCAAAATTCGAGAGTCTGTGCGGAGGTGAG 759
 234 ArgHisAsnArgSerLeuLysGluGlyValGlnArgAlaArgGluGluGluLys 253

760 AGACACAACAAGACTCTGAAGGAAGAGCGCTTCACGGGCGACGTGAGGAAGAGAGAA 819
 254 ArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGlu 273
 820 AGGAAGGAAATCACAAGGCATTTCCAGAGTAGTCCCTCAGCGACATCCAGGGCCAGATCGAG 879
 274 GlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeu 293
 880 CAGCAGACTGAGCGAAATATGAGCTCTGTGAGGAGAACACAGAGCTTCGAGAAAGCTG 939
 294 LysLysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAspLysValPheLysHis 313
 940 AAAAGCATCATCGATCATGATCAGCTCAGAGAGGAGCATCTGGACAAAATATTAAACAC 999
 314 LysAspLeuGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGlnMetLeuLys 333
 1000 AGAGAACTTCAGCAGAGAGCTGTGTGATGCCAAAGCTTGAGCAGGCCCAAGAAATGATGAAG 1059
 334 GluAlaGluArgHisGlnArgGluLysAspPheLeuLeuLysGluAlaValGluSer 353
 1060 GAAAGCGAGGAGCGACACAAACGAGAAAGAAATATTGCTGAACAGGCGAGAGGTGG 1119
 354 GlnArgMetCysGluLeuMetLysGlnGlnGlnThrHisLeuLysGlnGlnLeuAlaLeu 373
 1120 AAATCTCAGGCGAAGTGTGAGGAGCAAGAGACAGTCTCTGACAGGCTCAGCTCACTCTC 1179
 374 TyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluValPheThr 393
 1180 TACTCAGGAAGTTTGAAGAATTCAGAGCACACTAACTAAAGCAACAGAGGTGTTTGGC 1239
 394 ThrPheLysGlnGlnMetGluLysMetThrLysLysIleLysLysLeuGluLysGluThr 413
 1240 ACGTTCAACAGGAAATGACAAACCAACTAAGAAATGAAGAGCTGGAAAAAGGACACA 1299
 414 ThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeuGluMetAlaGluGlu 433
 1300 GCCCATGGAAGCCCGATTGAGAACTGTAAACAAGCTCTGTGGACATGATTGAAGAG 1359
 434 LysThrValArgAspLysGluLeuGluGluValLysIleGlnArgLeuGluLys 453
 1360 AAAGCACTGAGAGCTAAGAAATATGATGCTTTGTGATGAAATCGGAGGCTAGAGAAC 1419
 454 LeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgValGlnAspLeuSer 473
 1420 CTCTGCCGTGTTTACAGAGAGAGAAACGACTCCCAAAATAATCAGAGAGCCAGAA 1479
 474 AlaGlyGlnGlnLysSerLeuThrAspSerGlyProGluArgArgProGluGly----- 491
 1480 ATATCTGAAGAGGATGACCAAAAGTCAGCACAACTCCGATGAAGAGCCAGAGTCAAACGTC 1539
 491 ----- 491
 1540 TCTGTGATCAAGAGATTGACGAGAGAGGTTAATAGTGTCCAAACCCGCGTGAATAAT 1599
 491 ----- 491
 1600 CTGCCCACAGCTTCATGATAATTCATCATCCAGAGTCAACCCGCGACCAAGTCCAAAGAA 1659
 492 -----ProGlyAlaGlnAlaProSer 498
 1660 ACCCAACCCGAAATPAGCAGTCTTCAGAGAGAGTGTGACGCCGCTCTCAAGGAGCCAGAG 1719
 499 SerPro-----ArgValThrGluAlaProCysTyrProGlyAlaProSer 513
 1720 CAACCCCTCTGATCCCTTCACGGGATTCAGAGAGTCCCTGCTCCCTTAACCTCTCAG 1779
 514 ThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAla 528
 1780 GCTGAAGCCGAAGGAGGAGTGTGATGCTGAACCTCCCTCCCAAGGCC 1824

RESULT 28
AC110537/c

```

Center clone name: 44_A_4
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-Terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.950731
Consensus quality: 237099 bases at least Q40
Consensus quality: 239586 bases at least Q30
Consensus quality: 240394 bases at least Q20
Insert size: 210000; agarose-fp
Insert size: 240848; sum-of-coverage
Quality coverage: 12.5 in Q20 bases; agarose-fp
Quality coverage: 10.9 in Q20 bases; sum-of-coverage
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1 28687..contig of 28687 bp in length
* 28688 28787..gap of 100 bp
* 28788 30488..contig of 1701 bp in length
* 30489 30588..gap of 100 bp
* 30589 110109..contig of 79521 bp in length
* 110110 110209..gap of 100 bp
* 110210 113706..contig of 3497 bp in length
* 113707 113806..gap of 100 bp
* 113807 120823..contig of 7017 bp in length
* 120824 120923..gap of 100 bp
* 120924 133081..contig of 12158 bp in length
* 133082 133181..gap of 100 bp
* 133182 142866..contig of 9685 bp in length
* 142867 142966..gap of 100 bp
* 142967 147439..contig of 4473 bp in length
* 147440 147539..gap of 100 bp
* 147540 161221..contig of 13682 bp in length
* 161222 161321..gap of 100 bp
* 161322 183008..contig of 21687 bp in length
* 183009 183108..gap of 100 bp
* 183109 211297..contig of 28189 bp in length
* 211298 211397..gap of 100 bp
* 211398 239952..contig of 28555 bp in length
* 239953 240052..gap of 100 bp
* 240053 242048..contig of 1996 bp in length.
*
* Location/Qualifiers
*     1..242048
*         /organism="Mus musculus"
*         /mol_type="Genomic DNA"
*         /db_xref="taxon:10090"
*         /clone="RP23-44A4"
*         /clone_id="RPC1-23 Female Mouse BAC"
*
* Feature
*     1..28687
*         /note="assembly_fragment"
*         clone_end:SP6
*         vector_side:left
*
*     28788..30488
*         /note="assembly_fragment"
*
*     30589..110109
*         /note="assembly_fragment"
*
*     110210..113706
*         /note="assembly_fragment"
*
*     113807..120823
*         /note="assembly_fragment"
*
*     120924..133081
*         /note="assembly_fragment"
*
*     133182..142866
*         /note="assembly_fragment"
*
*     142967..147439
*         /note="assembly_fragment"
*
*     147540..161221

```

```

misc_feature      /note="assembly_fragment"
161322..183008_
misc_feature      /note="assembly_fragment"
183109..211297
misc_feature      /note="assembly_fragment"
211398..239952
misc_feature      /note="assembly_fragment"
240053..242048
misc_feature      /note="assembly_fragment"
clone_end:T7
vector_side:right

ORIGIN
Alignment Scores:
Pred. No.:      1.33e-51      Length:      242048
Score:          1260.00      Matches:      284
Percent Similarity: 61.90%      Conservative: 80
Best Local Similarity: 48.30%      Mismatches: 127
Query Match:      46.63%      Indels:      97
DB:              2          Gaps:      10

US-10-023-523-8 (1-530) x AC110537 (1-242048)
Qy      2 SerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAla 21
Db      87916 TCAGAACCCAGCACTCCCACTTCGCCAGGCGCAGCCCAACACACCCCAAAACCTTA 87857

Qy      22 AlaProAlaValGluAlaGluGlyProGlySerSerGluAlaProArgLysPro----- 39
Db      87856 GACCAGATTAAAGAGATATCTCATGATGATGGTAGAGGACATCAAGAGGACTTTAAT 87797

Qy      40 -----GluGlyAlaGlnAlaArgThrAlaGlnSer-----GlyAlaLeuArgAspVal 55
Db      87796 AACTCACTTAAAGAAATACAGGAGAACTCTGCTAAAGAGTTACAAGTCTTAAAGAAAA 87737

Qy      56 SerGluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysValAspIleAsn 75
Db      87736 CAGGAAACACAAATCAACACAGAGTAGAA-----GTCTTACAGAAAAAGAGGAAAAACATA 87683

Qy      76 GlnGlyGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluProGluAlaGlu 95
Db      87682 CAACCTGGTCATGGAATAGAACAAACCACTACTAGACCTTAAAGGGAAGTAGACACAAAT 87623

Qy      96 LysSerArgThrTyrValAlaArgAsnGlyGluProGluProThrProValValTyr--- 114
Db      87622 AAAGAAAACTCAAAGTAGGCGCAACACTGGAGATAGAAAACCCCTAGGAAAGAAATCTGGAAC 87563

Qy      115 -----GlyGluLysGluProSerLys 121
Db      87562 CATAGATTTGACATCAGCAACAGAAATACAGAGATGGAGGAGTACTGAGGAGGCTAGTGA 87503

Qy      121 sGlyAsp-----ProAsnThrGlu----- 127
Db      87502 GGGCGGACGGGGGACGGGTCTCCCTTTCTTCAGAGAGAGTACGATAGTTGGGGAG 87443

Qy      127 ----- 127
Db      87442 CAGGTGCGGGGTGTGTGTGACTGAAATGATCTGTTGAAATGGCAACATGGCAAGAGAC 87383

Qy      128 -----GluIleArgGlnSerAspGluValGlyAspAr 138
Db      87382 TAGAATTTGTGGTTAGGAGTAAACAGCACATGAGAGAAAGAGAAATCTCTGGACG 87323

Qy      138 GasPHisArgArg---ProGlnGlu-----LysLysLysAl 149
Db      87322 AGAAGCTGAACAGGTCCTCTCTGATGGCCAGCAGCAAGATTACAGAGTGCAGCAGCAACAAAGA 87263

Qy      149 aLysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrPr 169
Db      87262 GAAGACCTTAGGAAAGAGTTTGTACTGATGACGCGTTAAACACCTTTTCAACCCC 87203

Qy      169 oGluGluLysLeuAlaAlaLeuCysLysLysTyrAlaGluLeuLeuGluHisArgAs 189

```

RESULT 29
HSM807623
LOCUS

HSM807623

4647 bp

mrna

linear

PRI 30-AUG-2003

DEFINITION Homo sapiens mRNA; cDNA DKF2p451H216 (from clone DKF2p451H216).
 ACCESSION BX647477
 VERSION BX647477.1 GI:34366634
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 4647)
 AUTHORS Koehrer, K., Beyer, A., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
 Pobo, G., Han, M. and Wiemann, S.
 CONSRTM The German Human cDNA Consortium
 TITLE Direct Submission
 JOURNAL Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
 Neuberberg, GERMANY
 COMMENT
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by BMFZ (Biomedical Research Center at the
 Heinrich-Heine-University, Duesseeldorf/Germany) within the cDNA
 sequencing consortium of the German Genome Project. This clone
 (DKF2p451H216) is available at the RZPD in Berlin. Please contact
 the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
 information about the clone and the sequencing project is available
 at http://mips.gsf.de/proj/cDNA/.
 FEATURES
 Location/Qualifiers
 1..4647
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKF2p451H216"
 /tissue_type="human skeletal muscle"
 /clone_lib="451 (synonym: hicc1). Vector pSport1; host
 DH10B; sites NotI + SalI"
 /dev_stage="adult"
 4608..4613
 polyA_signal
 polyA_site
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,48e-53 Length: 4647
 Score: 1258.50 Matches: 275
 Percent Similarity: 61.18% Conservative: 89
 Best Local Similarity: 46.22% Mismatches: 136
 Query Match: 46.58% Indels: 95
 DB: 9 Gaps: 10
 US-10-023-523-8 (1-530) x HSM807623 (1-4647)
 Qy 18 ProSerGlnAlaAlaProAlaValGluAlaGlu----- 28
 Db 93 CCTTCTGATACCTCCCAAGATGAGGCTAATCACTCTGAACAGCTCTCAGCGGAACGA 152
 Qy 29 -----GlyProGly---SerSerGlnAlaPro-----ArgLysPro 39
 Db 153 CAGTCAACACCTCCCAAGGAGGACAGTTTCATCATACCAGTCAATGCGCTGGAGGAAGAA 212
 Qy 40 GluGlyAlaGlnAlaArgThr-----AlaGlnSerGlyAlaLeuArg 53
 Db 213 GATGCCAGGATCTCCACCCAGTCCACACCAGAGGAAGGCGAGTGTGCACCCC 272
 Qy 54 AspValSerGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysValAsp 73
 Db 273 GACATCTCTGAAGAGCTGAATCGACAGTGGGAAGACATCATTAACACTTAT-----GGG 326
 Qy 74 AsnAsnGlnGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluProGluAsp 93
 Db 327 TCTGCTGCCAGCAGCAGCGGAAGAGGCTCTGCCAGGGCCAGTGCAGCGCTGAGAAAT 386
 Qy 94 AlaGluLysSerArgThrTyrValAlaArgAsnGlnGluProGluProThrProValVal 113
 Db 387 GCAGAAATCACTGACAAC-----GAGGATGGGGAGCTGTGAGCAAAACAACACTGAAGAG 437

Qy 114 TyrGlyGluLysGluProSerLysGlyAspProAsnThrGluGluLeuArgGlnSerAsp 133
 Db 438 GCTGGAAGAGAACCCGTTGCTTCTGGAGAGCCACCCACTGCTCAA-----GAG 485
 Qy 134 GluValGlyAspArgAspHisArgArgProGlnGluLysLysAlaLysGlyLeuGly 153
 Db 486 CCCTGTCAGCTAAGGAGCAAAAA-----TTGGAAGAAGAAATCTTAAAGGATTAGGC 539
 Qy 154 LysGluLeuThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeu 173
 Db 540 AAGAAGCCCACTGCTAATGCAAAATCTGAACAAGTTGCAACACCGGAAGAAAAGTTT 599
 Qy 174 AlaAlaLeuCysLysLysTyrAlaGluLeuGluGluHisArgAsnSerGlnLysGln 193
 Db 600 GATTTTTTATTCAGAAGATGCTGAATTCGTGATGAACATCGTACTGAGCAAAAGAG 659
 Qy 194 MetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGlnLysAspHisLeuArgGly 213
 Db 660 TTAAGCTCTCCAAAAGAAACAGGTACAAATCAAAGAGAAAGACCCAGTTTACAAGGT 719
 Qy 214 GluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGln 233
 Db 720 GAACACAGCAGAGCTATCTCGCTCGAAGCAAAATGGAGAGTCTGTCCGGGAGCTGCAG 779
 Qy 234 ArgHisAsnArgSerLeuLysGluGlyValGlnArgAlaArgGluGluGluLys 253
 Db 780 AGACACAAACAGACTCTGAAGGAAGAGGGCGCTTCAGCGGCACGTCAGAGAAAGAGAA 839
 Qy 254 ArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGlu 273
 Db 840 AGGAAGAAATCAACAGCCATTTCAGAGTAGTACCTCCAGACATCCAGGGCCAGATCGAG 899
 Qy 274 GluHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeu 293
 Db 900 CAGCAGAGTGCAGCAAAATGAAGCTCTGTGAGAGAACACAGAGCTTGCAGAAAGCTG 959
 Qy 294 LysLysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAspLysValPheLysHis 313
 Db 960 AAAAGCACCATCATGATGATGAGTCAGAGAGGAGCATCTGCACAAAATATTTAAACAC 1019
 Qy 314 LysAspLeuGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnLeuMetLys 333
 Db 1020 AGAGAACTGCAGCAGAGAGCTGGTGGATGCAAGCTTGAGCAGGCCCAAGAAATGATGAAG 1079
 Qy 334 GluAlaGluGluArgHisGlnArgGluLysAspPheLeuLeuLysGlnAlaValGluSer 353
 Db 1080 GRAGCGAGGAGCGACACAAACGGAAGGAATATTTGCTGAACCCAGCAGCAGAGTGG 1139
 Qy 354 GlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeu 373
 Db 1140 AAATTCAGCGGAAGTGTGGAAGGACAGACAGTCTCTGAGGCTCAGCTCACTCTC 1199
 Qy 374 TyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluValPheThr 393
 Db 1200 TACTCAGGAAGTTTGAAGAAATTCAGAGCACAATCACTAAAGCAGAGAGTGTGTTGCC 1259
 Qy 394 ThrPheLysGlnGlnMetGluLysMetThrLysLysLysLysLysLysGluThr 413
 Db 1260 ACGTTCAACAGGAATGGACAAAACAACTAAGAAAATGAAGAGCTGGAAGAAAGGACACA 1319
 Qy 414 ThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeuGlnMetAlaGluGlu 433
 Db 1320 GCCACATGGAAAGCCCGATTTGAGAACTGTAAACAAGCTCTGTGTGACATGATTAGGAG 1379
 Qy 434 LysThrValArgAspLysGluLeuGluGlnValLysIleGlnArgLeuGluLys 453
 Db 1380 AAGCACTGAGAGCTAAAGAAATATGAGTGTGTTGTGATGAATAATCGGAGGCTAGAGAAC 1439
 Qy 454 LeuCysArgAlaLeuGlnThrGluArgAsnAspLysAsnLysArgValGlnAspLeuSer 473
 Db 1440 CTCTGCCGTCTTTACAAAGAGAGAAACAACTCCACAAAAAATACAGAGACGCGAGAA 1499
 Qy 474 AlaglyGlyGlnGlySerLeuThrAspSerGlyProGluArgArgProGluGly----- 491

```

Db 1500 ATATCTGAAGAGGATGACCAAGTTCAGCAACTCCGATGAGAGCCAGATCAAAAGTC 1559
QY 491 -----
Db 1560 TCTGTGGATCAAGAGATTGACGACAGAGGAGGTTAATAGTGTCCAAACCGCGTGAAAAAT 1619
QY 491 -----
Db 1620 CTGGCCACAGCCTTCATGATTAATCATCATCAGAGTCAACCCCGCACAGTCCAAAGAA 1679
QY 492 -----
Db 1680 ACCCAACCCGAATAGGAGTTCTCAGGAGAGTGCTGACCGCGCTCTCAAGGAGCCAGAG 1739
QY 499 SerPro-----ArgValThrGluAlaProCysTyrProGlyAlaProSer 513
Db 1740 CAACCCCTCTGATCCCTTACGGGATTCAGAGAGTCCCTGCTCCCTAACTCCTCAG 1799
QY 514 ThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAla 528
Db 1800 GCTGAGCCGAGGAGGAGTGTGCTGAACTCTCTCCAGGCC 1844

RESULT 30
D89999
LOCUS      4704 bp mRNA linear VRT 26-APR-2000
DEFINITION Gallus gallus MDP77 mRNA for muscle derived protein, complete cds.
ACCESSION D89999
VERSION   D89999.1 GI:7619883
KEYWORDS  Gallus gallus (chicken)
SOURCE    Gallus gallus
ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.

REFERENCE
  1 (sites)
    Uyeda, A., Fukui, I., Fujimori, K., Kiyosue, K., Nishimune, H., Kasai, M.
    and Taguchi, T.
    MDP77: A novel neurite-outgrowth-promoting protein predominantly
    expressed in chick muscles
    Biochem. Biophys. Res. Commun. 269 (2), 564-569 (2000)
    20175243
    10708594
    Uyeda, A.
    Direct Submission
    Submitted (24-DEC-1996) Atsuko Uyeda, Osaka National Research
    Institute, Department of Organic Materials, 1-8-31, Midorigaoka,
    Ikeda, Osaka 563-8577, Japan (E-mail: au-uyeda@onri.go.jp,
    Tel: +81-727-51-9524, Fax: +81-727-51-9628)
    Location/Qualifiers
      1..4704
        /organism="Gallus gallus"
        /mol_type="mRNA"
        /db_xref="taxon:9031"
        /sex="male"
        /tissue_type="muscle"
        /dev_stage="postnatal 8 days"
      1..4704
        /gene="MDP77"
        /gene="MDP77"
        /function="neurite-outgrowth-promoting protein"
        /note="old name: muscle derived factor F460
        with molecular mass 77 k"
        /codon_start=1
        /product="muscle derived protein"
        /protein_id="BA094755.1"
        /db_xref="GI:7619884"
        /translation="MENDQFTKQCOVTTSPDNDQSGKAEPPVPSQPLSPTNQTSA
        QEMATCDISELNLQLEDI IKYGSASALIVEKGTAAETDKPEKQVSGMEDACECD
        VNESEKQKPAFGDRAKPEASKEQKRLKLGKREAILLQSLNKLITTEFKL
        DLLFKYAELEBEHFAEQKLYLQKRAQITKEKQDQSEHRAILARSKLESICRE
  
```

```

LOHNKTLKEETIORAREBEKKEITNFQGTLSIEIOACIEBOQGERNMKLCQENTLE
AEKLSIIDQVELREBEHLKIKHRELOQKLVDAKLEQSCNMKKEAEREHOREKEVLL
NQAEKLDQAKMLKEQETVLAQITLYSERFEFKTLTKSNEVATFKQMEKWTKK
MKLEKDTATWKSFRFENCRLAQDITTEERAMTKKEYECFLKIQRLNLCRALQGBRN
SHYKTKIQALPEEVNGDILDEDDANTNPSSEQASIELCAAGNMQLQELAEAPRV
EUKAETLPSDGSNPETCNVQMECAISVPELPSHLTSQPEAGNHCEQFSMTSAPPEH
MPAATENMTLTIENMPKTPKSMPEMPVPTPTESVPIPTGVPTPKIMPATPSVP
TLQNQTSAPLIGNMPASTKSTPKAVEHVDDIAELFIPDQPAEQKGDITDMEAVD"

ORIGIN
Alignment Scores:
Pred. No.:      2,05e-52      Length:      4704
Score:          1240.00      Matches:      271
Percent Similarity: 65.87%      Conservative: 88
Best Local Similarity: 49.72%      Mismatches: 131
Query Match:      45.89%      Indels:      55
DB:              5          Gaps:      11

US-10-023-523-8 (1-530) x D89999 (1-4704)
QY 2 SerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAla 21
Db 238 ACCTCACCTACGACGACACACACAGGG-----CAAGCAGGCGGAACT 282
QY 22 AlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGly 41
Db 283 GTCCCGCTCTCACAGCGCCCTTTCCGCC--ACAAACCAACACAGTCCCGAGCTGAGTG 339
QY 42 AlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArg 61
Db 340 GCCACGTGT-----GACATCTCAGAGGAGCTGAACCCG 372
QY 62 GlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGluGlu 81
Db 373 CAACTGGAGGACATTATTAACATAT-----GGGTCT 405
QY 82 AspGlyAla-----GlnGlyGluProAlaGluProGluAspAlaGluLysSer 97
Db 406 GCAGCGAGTCTGTAGAGAGGAGGAGCCTACAGCAGAACTACAGCAGAGGAGAA 465
QY 98 ArgThrTyrValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGluLys 117
Db 466 GATGTGGCAGTATGGAGGATGCAGAGTGTGAGATGTAATCAAGAAAGTGAGAAAGAC 525
QY 118 GluProSerLysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAsp 137
Db 526 AAACCCAGCTCTCGAGATGCTTCA-----AGAGCAAGAGGAGCCAGTGCCAGC 573
QY 138 ArgAspHisArgArgProGlnGluLysLysAlaLysGlyLeuGlyLysGluIleThr 157
Db 574 AAGAACAAAG-----CTGGAGAGAGAAATCTCTAAGGACTAGGAGAGGAGCCACC 627
QY 158 LeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCys 177
Db 628 CTCCTCATGCAAGACTTGAACAGCTGACTACTCTCAGAGGAGAGAGCTGTGACTTT 687
QY 178 LysLysTyrAlaGluLeuLeuGluHisArgAsnSerGlnLysGlnMetLysLeuLeu 197
Db 688 AAGAAGTATCTGAGTGTCTTGAGGAGATCTGTGCTGACGAGAGAGAGCTCAAGTACCTG 747
QY 198 GlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLys 217
Db 748 CAGAGAGGCGAGCCGCCATCACCAGAGAGAGAGCAGTTCGACAGAGTGCAGCAGCCGA 807
QY 218 AlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArg 237
Db 808 GCATCTCTGCTCGACGAGCTTGTAGAGCTGTGCGGGAGCTGCCAGAGGAGCAGCAAAA 867
QY 238 SerLeuLysGluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluVal 257
Db 868 ACCCTCAGAGGAGAAACAATTTCAGCGGCGACGCGAGAGAGTGAAGAGAGGAGAAATA 927
QY 258 ThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGlu 277
  
```



```

Db 928 ACAATCATTTCCAGGCGACGCTGAGTGAATCCAGGCTCAGATTGAGCAGCAAAAGTGAG 987
Qy 278 ArgAsnSerLysLeuArgGlnGlnLeuMetGluLeuAlaGluArgLeuLysLeuLeu 297
Db 988 AGGAACATGAAGCTCTGCCAGGAGAACACAGAGCTGGCAGAGAAGCTGAAAGCAATTATT 1047
Qy 298 GluGlnTyrGluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGln 317
Db 1048 GACCAGTATGAGCTGGGGAGAGCACCCTTGACAAATATTTAAGCACAGAACTTCAA 1107
Qy 318 GlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluGlu 337
Db 1108 CAGAAATTTGGTATGCCAAGTTGGAGCAGTCTCAGGAAATGATGAAGAGCAGAGG 1167
Qy 338 ArgHisGlnArgLysAspPheLeuLysGluAlaValGlnSerGlnArgMetCys 357
Db 1168 CGACATCAGAGGAAGGAATATCTCTGAATCAAGCCGAGATGGAGGTACAGGCC 1227
Qy 358 GluLeuMetLysGlnGlnGlnHisLysLysGlnGlnLeuAlaLeuTyrThrGluLys 377
Db 1228 AAAATGTTAAAGGAACAGGAGACAGTCTCTGAGGACAGATCACTCTCTACTCTGAGAGA 1287
Qy 378 PheGluGlnPheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGln 397
Db 1288 TTTGAAGAATTTTCAAGAAACATTTGACCAAAAGCAATGAAGTGTCTGCTACCTTCAAAAG 1347
Qy 398 GluMetGluLysMetThrLysLysIleLysLysLeuGluLysGluThrThrMetTyrArg 417
Db 1348 GAGATGGAGAAATGACAAAGAAATGAAGAAGTTGGAAGAGGATCTCTACATGTGAAA 1407
Qy 418 SerArgTyrGluSerSerSerLysLysAlaLeuLeuGluMetAlaGluLysThrValArg 437
Db 1408 TCCAGGTTTGAAGAACTGTAACAGACATTACTGACATGATTGAAGAGAAAGCCATGAGG 1467
Qy 438 AspLysGluLeuGluGluGlnValLysIleGlnArgLeuGluLysLeuCysArgAla 457
Db 1468 ACCAAGAAATACAGAGTCTTTGCTGTAATATCCAGAGGCTAGAACCTTTGCCGAGCT 1527
Qy 458 LeuGlnThrGluArgAsnAspLeuAsnLysArgValGlnAspLeu-----Ser 473
Db 1528 CTCAGAGAGAGGAATGAATGTTACAGAAAAATAAAACAGCAGCTCCTCTGAAGAG 1587
Qy 474 AlaGlyGlnGlySerLeu---ThrAspSerGlyProGluArgArgProGluGly-Pr 492
Db 1588 GTGAATGGAATGACATCTTAGAAGAGAGCAGCATGCCAATACAAACCTTCTCTCTCC 1647
Qy 492 oGlyAlaGlnAlaProSerSerProArgValThrGluAlaProCys----- 507
Db 1648 GAGCAGGCAAGCATTGAGCTATGCTGCTGACAGAACATGCTGCAGGAGCTAGCTGAA 1707
Qy 508 -----TyrProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnG 524
Db 1708 GCTTTCAGGGTATCCC-----ACAAAGCAGAGAGACCCCTCCCA 1746
Qy 524 uProThrSerAla 528
Db 1747 AGCAGCGGACGA 1759

```

RESULT 31

```

E26533
LOCUS E26533 4723 bp DNA linear PAT 18-JUN-2001
DEFINITION Protein with neurite outgrowth activity.
ACCESSION E26533
VERSION E26533.1 GI:13026200
KEYWORDS JP 1999147897-A/1.
SOURCE unidentified
ORGANISM Takahisa,T
REFERENCE 1 (bases 1 to 4723)
AUTHORS Takahisa,T
TITLE Protein with neurite outgrowth activity
JOURNAL Patent: JP 1999147897-A 1 02-JUN-1999;

```

AGENCY OF IND SCIENCE & TECHNOL

```

OS Unidentified
PN JP 1999147897-A/1
PD 02-JUN-1999
PF 13-NOV-1997 JP 1997331242
PR TAKAHISA TAGUCHI
PI C07K14/52,A61K38/00,A61K38/00,A61K38/00,C12N15/09,A61K37/02,
PC A61K37/02,
PC A61K37/02,C12N15/00
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..4723
FT /organism='Unidentified'.

```

FEATURES

```

source 1..4723
/organism='unidentified'
/mol_type='genomic DNA'
/db_xref='taxon:32644'

```

ORIGIN

Alignment Scores:

```

Pred. No.: 2,06e-52 Length: 4723
Score: 1240.00 Matches: 271
Percent Similarity: 65.87% Conservative: 88
Best Local Similarity: 49.72% Mismatches: 131
Query Match: 45.89% Indels: 55
DB: 6 Gaps: 11

```

US-10-023-523-8 (1-530) x E26533 (1-4723)

```

Qy 2 SerSerProGlyGlnProGluAlaGlyProGluAlaGlnGluArgProSerGlnAla 21
Db 238 ACCTCACCTACGACGAGCAACACGCGG-----CAAGCAAGCGAGAACCT 282
Qy 22 AlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGly 41
Db 283 GTCCCGCTCTCACAGCCCTTTCCGCC-----ACAACCAACACGAGTCCCGCTGAGATG 339
Qy 42 AlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArg 61
Db 340 GCCACGCTG-----GACATCTCAGAGGAGCTGAACCGG 372
Qy 62 GlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGlyGlu 81
Db 373 CAACTGGAGGACATATTAAACCATAT-----GGGTCT 405
Qy 82 AspGlyAla-----GlnGlyGluProAlaGluProGluAspAlaGluLysSer 97
Db 406 GCAGCGAGCTGGTAGAAGAGGAGGCACTACAGCAGAAACTGACAGCCAGAGAAGAA 465
Qy 98 ArgThrTyrValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGluLys 117
Db 466 GATGTGGCAGTATGGAGGTGCGAGGTGTGAGGATGTAATGAAGAAAGTGAAGAGAC 525
Qy 118 GluProSerLysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAsp 137
Db 526 AAACCAGCTCTCTGGAGATGCTTCA-----AGACAAAGGAGCCCGAGTCCGACG 573
Qy 138 ArgAspHisArgArgProGlnGluLysLysLysAlaLysGlyLeuGlyLysGluLeuThr 157
Db 574 AAGGAACAAAAG-----CTGGAGAAGAAAATCTCTGAAGAGCTAGGGAAGAGACCCACC 627
Qy 158 LeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCys 177
Db 628 CTCCTCATGCAAGCTTGAACAGCTGACTACTCTCAGAGGAGAGAGTGGACCTGTTATT 687
Qy 178 LysLysTyrAlaGluLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeu 197
Db 688 AAGAAAGTATGCTGAGTTCTTGGAGGAGCATCGTGTGAGCAGAGCATCAAGTACCTG 747
Qy 198 GlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLys 217

```


QY 38 LysProGluGluAlaGlnAlaArgThrAlaGlnSerGlyAlaLeu----- 52
 Db TACTCAACCTCAGGCCAAGCG---CCAGAGCAAGAGGAGGCTGCATCCCGAGAGGGA 318
 QY 53 ---A-gAspValSerGluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCys 71
 Db GCCCATGATGTCGCGAAGAGTTGAGCAGCACTGGAAGACATCATAGCAGCTAT--- 375
 QY 72 ValAspAsnAsnGlnGlyGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluPro 91
 Db -----GGTCTGCTGCCAGTCCC 393
 QY 92 GluAspAlaGluLysSerArgThrTyrValAlaAlaArgAsnGlyGluProGluProThrPro 111
 Db GAGGAGGAGAGCGCTCTCAACT----- 417
 QY 112 ValValTyrGlyGluLysGluProSerLysGlyAspProAsnThrGluGluLeu----- 129
 Db -----AAGGAGCAGCGCCCAACACAGAGGCGACCAAGCAATGAGGATGTTGACTAT 468
 QY 130 A-gGlnSerAspGluValGlyAspArgAspHisArgArgProGln----- 144
 Db -----GluLysLysLysAlaLysGlyLeuGly 153
 QY 145 -----GluLysLysLysAlaLysGlyLeuGly 153
 Db AAGAGCGCTGTCAGCAATAAAGAGCAAAAGCTGGAAGAGAAATCTTAAAGAGATAGGT 588
 QY 154 LysGluLeuThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeu 173
 Db AARGAGCTTAACCTTGTCTATGTCGAAACCTTGACAGCTGCAAGCCTGGAAGAAAGCTT 648
 QY 174 AlaAlaLeuCysLysLysTyrAlaGluLeuLeuGluGluHisArgAsnSerGlnLysGln 193
 Db GATTTTTTATTCAGAGAGTATATCTGAATGCTGATGTAACATCGCACCGACCAAGAGAA 708
 QY 194 MetLysLeuLeuLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGly 213
 Db CTGAAGCTCTACTGAAACAGAGCGGCGAGACCCAGAGAGAGAGGACCGAGTTCAGAGT 768
 QY 214 GluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGln 233
 Db GAGCAACACAGAGCGCTCTGCTCGAAGCAAACTCGAGAGTCTGTGTCGGAGCTACAG 828
 QY 234 ArgHisAsnArgSerLeuLysGluGlyValGlnArgAlaArgGluGluGluLys 253
 Db AGGCAACAAACAGAGCGTGAAGAGAAACCTCCACCGCGCGCGAGAGAGAGAGAG 888
 QY 254 ArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGlu 273
 Db AGGAAGAGATCAACAGTCACTTCAAACTACCTGCAGATATCCAGACTCAGATCGAA 948
 QY 274 GlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeu 293
 Db CAACAGAGTGAAGCAATATGAAGCTCTGCGAGAGAAACACAGAGCTTGACAGAAACTG 1008
 QY 294 LysLysLeuLeuGluGlnTyrGluLeuArgGluGluHisLysLysValPheLysHis 313
 Db AAGAGCATCATTCACCAAGTACGAGCTCAGAGAGAGCACTGCAACAAATATTCAAACAT 1068
 QY 314 LysAspLeuGlnGlnGlnLeuValAlaLysLeuGlnGlnAlaGlnGluMetLeuLys 333
 Db CGAGAACTGCGAGAGAGAGTGGTGGATGCGAACTTGAAGAGCTCAAGAACTGATGCGAG 1128
 QY 334 GluAlaGluArgHisGlnArgGluLysAspPheLeuLeuLysGluAlaValGluSer 353
 Db GAGCAGAGGAGCGGCACAGCGGAGAGGAATATTGCTGAACCAAGCGCGGAGAGTGG 1188
 QY 354 GlnArgMetCysGluLeuMetLysGlnGlnThrHisLeuLysGlnGlnLeuAlaLeu 373
 Db AACTTCAGGCCAAAGTGTGAGGAGCAGAGAGCGGTGTCGAGGCTCAGCTCACTCTC 1248

QY 374 TyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluValPheThr 393
 Db TACTCAGAAAGGTTTGAAGGATTCAGAGTACACTGACAAAAAGCAATGAAGTGTTCGCC 1308
 QY 394 ThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLysLysLysLysGluThr 413
 Db ACTTTTCAAGCAGGAATGATAGACAACTTACAAATGAAGAGCTGGAGAGGACACC 1368
 QY 414 ThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeuGluMetAlaGluGlu 433
 Db GCCACGTGGAAGCGCGATTTGAGAAATTGAAACAAAGCCCTGTGGACATCATCGAAGAG 1428
 QY 434 LysThrValArgAspLysGluLeuGluGlyLeuGlnValLysLysLysLysLysGluLys 453
 Db AAGCCCTGCGAGCTAAAGAAATACGAGTGTTCGTGATGAATTCAAAGCTCGAAGAAC 1488
 QY 454 LeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgValGlnAspLeuSer 473
 Db CTGTGTCGTGCTTTTACAAGAGGAGAGAAAGCACTCTACAAGAAATACAGAGAACAAAA 1548
 QY 474 AlaGlyGlnGlnLysLeuThrAspSerGlyProGluArgArgProGluGlyProGly 493
 Db ATGTCCGAAAGAGGAGACCAAGGTTTCAGGCACCTCTGAGGAAGAGCCAGAG---CCAAGC 1605
 QY 494 AlaGlnAlaProSerSerProArgValThrGluAlaProCysTyrProGlyAlaProSer 513
 Db GTCTCTGAAATGAAGAGGTGAGCGCAGAGAGCAACAGTTTTCAAAAGCTGTGGAA 1665
 QY 514 ThrGluAlaSerGlyGlnThr-----GlyProGlnGluProThr 526
 Db AACCTGCGCCAGAGCTTTTACCATCTTCATCATCCAGAGTTCACCCAGACCAACCAACG 1725

RESULT 33
 AF422245 4394 bp mRNA linear ROD 24-JAN-2002
 LOCUS Mus musculus muscle-derived protein MDP77 variant 2 (Mdp77) mRNA,
 DEFINITION complete cds.
 ACCESSION AF422245
 VERSION AF422245.1 GI:17026375
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 4394)
 AUTHORS Benson, K.F. and Chada, K.
 TITLE Molecular characterization of the mouse In(10)17Rk inversion and
 identification of a novel muscle-specific gene at the proximal
 breakpoint
 JOURNAL Genetics 160 (1), 279-287 (2002)
 MEDLINE 21663722
 PUBMED 11805063
 REFERENCE 2 (bases 1 to 4394)
 AUTHORS Benson, K.F.
 TITLE Direct Submission
 JOURNAL Submitted (21-SEP-2001) Medical Genetics, University of Washington,
 1705 Pacific Street NE, Seattle, WA 98195, USA
 FEATURES
 location/Qualifiers
 1..4394
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /chromosome="10"
 /map="between Cni-rs2 and Mtap7"
 /sex="female"
 /tissue_type="heart"
 /dev_stage="adult"
 1..4394
 /gene="Mdp77"
 157..2211
 /gene="Mdp77"
 /note="muscle-specific; similar to Gallus gallus"

gene
 CDS

muscle-derived protein encoded by GenBank Accession Number D89999"
 /codon_start=1
 /product="muscle-derived protein MDP77 variant 2"
 /protein_id="AA13910.1"
 /db_xref="GI:17026376"
 /translation="MEINHPDQLSVRHPTPPGSSINONGPKGKODGRCSTSGOAP
 QEGSLHEKGAHVAEELSQLEDIISTYGAASPEGDSASETKGPPNTEAPDNEDV
 DEFTTSEIDREPTAPPEAPAAKEPVSNKEQLEKILKLGKKEANLLMKNLNUQAP
 EEKDLFFKFKYTBELLDEHTEQKKLLKQQAOTQREKQDQSENRVAVLARSLES
 LCRLDFQNTKLTETLQAREEERKEITSHFTQTLTDIQIQQOESERNMKLCOE
 NTELAELKSIIDQYELREHLDKIFKHLRELQKLVDALEAEALMQEAERHREK
 EYLINQAAKWLQAKVLKEQTVLQALTYLSEFEFSTLTLSNEVATFKOEMDK
 TTKKMKLEKDTAKWKEFNCNKALLDMEKALRAKSECFWKIQLLENLCRALQ
 BEKELYKLRKAKSEKEQVORTSEEPSEPSVSENEVEDAEANSFOKAVENLATA
 FTILHPEFTPDQTERQLAVQPSGSDVTHORPETARLNHPSPADSGSPRPVPA
 QVAEGVCEATPAPTACTPAEAELOSQGLPAENTPGPKHPEANTSQAPLSPAQG
 SLSVYEAKYDISPSESGDSAVVPGESREQPPPEVTDIPVGPSTGLPREPDA
 CLNG
 V"

ORIGIN

Alignment Scores:

Pred. No.: 6,28e-52 Length: 4394
 Score: 129.50 Matches: 265
 Percent Similarity: 63.89% Conservative: 80
 Best Local Similarity: 49.07% Mismatches: 132
 Query Match: 45.50% Indels: 63
 DB: 10 Gaps: 8

US-10-023-523-8 (1-530) x AF422245 (1-4394)

Qy 18 ProSerGlnAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArg 37
 Db 202 CCCCAGGGGACAGCTCATCTCAATCAAAACGGCCGGGAAGCAAGATGCGAGCGG 261
 Qy 38 LysProGluGlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeu----- 52
 Db 262 TGCTCAACCTCAGGCCAAGCG---CCAGAGCAAGAGGGAAGCCTGTCATCCGAGAGGGA 318
 Qy 53 ---ArgAspValSerGluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyCys 71
 Db 319 GCCCATGATCTCGGAGAGTTGACAGGCACTGGAAGACATCATAGACGATAT--- 375
 Qy 72 ValAspAsnAsnGlnGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluPro 91
 Db 376 -----GGGTCCTGCTGCCAGTCCC 393
 Qy 92 GluAspAlaGluLysSerArgThrTyValAlaArgAsnGlyGluProGluProThrPro 111
 Db 394 GAGGAGGAGACCCCTCTGAACT----- 417
 Qy 112 ValValTyGlyGluLysGluProSerLysGlyAspProAsnThrGluGluIle----- 129
 Db 418 -----AAGGAGCAGCCCCCAACACAGAGGACCCAGCAATGAGGATGTGACTAT 468
 Qy 130 ArgGlnSerAspGluValGlyAspArgAspHisArgArgProGln----- 144
 Db 469 GAAGAAACCACTGAAGATAGACAGAGAACCCACCTGCTCCGAAAGAGCCAGCCGAGCC 528
 Qy 145 -----GluLysLysLysAlaLysGlyLeuGly 153
 Db 529 AAAGAGCCTGTGACAAATAAGAGCAAAAGCTGGAAAGAAATCTTAAAGATTAGT 588
 Qy 154 LysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeu 173
 Db 589 AAAGAAAGCTAACTTGCTCATGCAAAACCTGAACAAGCTCAAGACCTGAAGAAAGCTT 648
 Qy 174 AlaAlaLeuCysLysLysTyValAlaGluLeuGluGluHisArgAsnSerGlnLysGln 193
 Db 649 GATTTTTTATCAAGAGATATACTGATGCTGGATGACATCGCACCGAGCAAGAA 708
 Qy 194 MetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGly 213
 Db 194 MetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGly 213

Db 709 CTGAAGCTCCTACTGAAACACAGAGCGCGCAGACCCAGAGAGAGAGACCACTGTGCAGAGT 768
 Qy 214 GluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLysLeuGluGln 233
 Db 769 GAGCACACAGAGCGCTGCTTGTCTGAAAGCAAACTCGAGAGCTGTGTGCGGAGCTACAG 828
 Qy 234 ArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGluGluGluGluLys 253
 Db 829 AGGCACAACAGAGCCTGAAGAAACACCTTCAACCGGCGCGAGAGAGAGAGAGAG 888
 Qy 254 ArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGlu 273
 Db 889 AGGAAGAGATCACAGATCACTTTCAACTACCTTGACAGATATCCAGACTCAGATCGAA 948
 Qy 274 GlnHisAsnGluValArgAsnSerLysLeuArgGlnGlnAsnMetGluLeuAlaGluArgLeu 293
 Db 949 CAACAGAGTGAAGCAATATGAAGCTCTGCAGAGAGAACACAGAGCTTGAGAGAAACTG 1008
 Qy 294 LysLysLeuIleGluGlnTyArgGluArgGluGluHisIleAspLysValPheLysHis 313
 Db 1009 AAAGCATCATGACAGTACAGCTCAGAGAGAGCATCTGGACAAATATTCAAACAT 1068
 Qy 314 LysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGlnMetLysLys 333
 Db 1069 CGAGAACTGACAGCAGAGCTGCTGATCGCAAACTTGAAGAGGCTCAAGAACTGATCAG 1128
 Qy 334 GluAlaGluGluArgHisGlnArgGluLysAspPheLeuLeuLysGluAlaValGluSer 353
 Db 1129 GAGGAGAGGAGCGCACAGCGGAGAGAAATATTGTGTAACACGCGCGAGAGTGG 1188
 Qy 354 GlnArgMetCysGluLeuMetLysGlnGlnGlnThrHisLeuLysGlnGlnLeuAlaLeu 373
 Db 1189 AAATTCAGGCCAAAGTGTCTGAGAGCAGAGAGCGGTGTGCGAGGCTCAGCTCCTC 1248
 Qy 374 TyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluValPheThr 393
 Db 1249 TACTCAGAGAGGTTGAGGAGTTCAGAGTACCTGACAAAGAAAGCAATGAAGTGTTC 1308
 Qy 394 ThrPheLysGlnGluMetGluLysMetThrLysLysIleLysLysLeuGluLysGluThr 413
 Db 1309 ACTTTCAGCAGGAAATGGATAGACAACTAAGAAATGAAGAGCTGGAGAGGACACC 1368
 Qy 414 ThrMetTyArgSerArgTrpGluSerSerAsnLysAlaLeuLeuGluMetAlaGluGlu 433
 Db 1369 GCCACGTGGAAGCCCGATTGAGATTGTAACAAAGCCCTGTGTGGACATGATCGAAGAG 1428
 Qy 434 LysThrValArgAspLysGluLeuGlnValLysLysIleGlnArgLeuGluLys 453
 Db 1429 AAAGCCTGCGAGCTAAAGATACAGTGTTCGTGATGAATTCGAAAGCTGGAGAAC 1488
 Qy 454 LeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgValGlnAspLeuSer 473
 Db 1489 CTGTGTGCTGTTTACAGAGGAGAGAGAAAGAACTCTACAGAAATTCAGAGAGCAAAA 1548
 Qy 474 AlaGlyGlnGlnSerLeuThrAspSerGlyProGluArgArgProGluGlyProGly 493
 Db 1549 ATGTCGGAAGAGGAGACACAGGTTACCGGACCTCTGAGGAGAGCCAGAG---CCAAGC 1605
 Qy 494 AlaGlnAlaProSerSerProArgValThrGluAlaProCysTyProGlyAlaProSer 513
 Db 1606 GTCTCTCAAAATGAAGAGGTGGAGCGAGAGAGCAACAGTTCCTCAAAAGCTGTGGA 1665
 Qy 514 ThrGluAlaSerGlyGlnThr-----GlyProGlnGluProThr 526
 Db 1666 AACCTGCCACAGCTTTTACCATCTTCATCTCCAGAGTTCACCCAGACCAACCAACG 1725

RESULT 34

BC059099

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

BC059099 1226 bp mRNA linear ROD 20-OCT-2003
 Mus musculus cDNA clone IMAGE:4947346, partial cds.

BC059099.1 GI:37748168

SOURCE
ORGANISM Mus musculus (house mouse)
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (Bases 1 to 1226)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uedin, R.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Viall, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE 22388257
PUBMED 12477932
REFERENCE 2 (Bases 1 to 1226)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (01-OCT-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
 Email: cgaps-remail.nih.gov
 cDNA Procurement: Jeffrey Green M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
info@bcgc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letitia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalios, Michael Smith, Lorraine Spence, Jeff Scott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 131 Row: 1 Column: 12
 This clone was selected for full length sequencing because it passed the following selection criteria: GenomeScan gene prediction.
FEATURES
source
 1. 1226
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone IMAGE:4947346
 /tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating ductal carcinoma. 5 month old virgin mouse."
 /clone_lib="NCI CGAP Mam6"
 /lab_host="DH103"
 /note="vector: pCMV-SPORT6"

Alignment Scores:
 Pred. No.: 7,34e-44 Length: 1226
 Score: 1054.00 Matches: 217
 Percent Similarity: 68.43% Conservative: 54
 Best Local Similarity: 54.80% Mismatches: 63
 Query Match: 39.01% Indels: 62
 DB: 10 Gaps: 4
 US-10-023-523-8 (1-530) x BC059099 (1-1226)
 QY 19 SerGlnAlaAlaProAlaValGluAlaGluGlyPro-----GlySerSerGlnAlaPro 36
 Db 216 TCTCAAGCAAAATGATATTCTTCAACATCAAGCAAGCCAGTGTGGTGGCACTAAGAAA 275
 QY 37 ArgLysProGluGlyAlaGlnAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 56
 Db 276 CATTCACTGGAAGGGATGAAGC-----ACT 302
 QY 57 GluGluLeuSerArgGlnLeuGluAspLleLeuSerThrTyrcysVal---AspAsn 75
 Db 303 GACTTTATTACAAAGAACAGAAATTTGGTGAGCTCAGTATTCTGTACACAGAGAAAAGA 362
 QY 76 GlnGlyGlyProGluGluAspGlyAlaGlnGlyGluProAlaGluProGluAlaGlu 95
 Db 363 GAAGAATTCCTGGACGAGAGCTCGAACAGGTCTCTCTGATGCCAGCAAGATTCAGAG 422
 QY 96 LysSerArgThrTyrcysValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 115
 Db 423 TGCAGCAGG----- 431
 QY 116 GluLysGluProSerLysGlyAspProAsnThrGluGluLeuArgGlnSerAspGluVal 135
 Db 431 ----- 431
 QY 136 GlyAspArgAspHisArgArgProGlnGluLysLysLysLysLysLysLysLysLys 155
 Db 432 -----AACAAAGAGAGACCTTAGGAAAAAGAA 458
 QY 156 IleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaAla 175
 Db 459 GTTTTATTACTGACAGCGGTAAACACCCCTTTCAACCCAGAGAGAGAGTGCAGCT 518
 QY 176 LeuLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 195
 Db 519 CTCTGTGAAGATATGCTGATCTCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 578
 QY 196 LeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHis 215
 Db 579 ATTCTGCAG 638
 QY 216 SerLysAlaValLeuAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 235
 Db 639 AGCAAGCCATCTTGGCAAG 698
 QY 236 AsnArgSerLeuLysGluGluGlyValGlnArgAlaAlaAlaAlaAlaAlaAlaAlaAla 255
 Db 699 AATAAGACCTTAAAG 758
 QY 256 GluValThrSerHisPheGlnValThrLeuAsnAspLleGlnLeuMetGlnHis 275
 Db 759 GAAGCAACAGACCATTTCCAGATAACTTAAATGAATCAAGCTCAGTTGAAACAACAT 818
 QY 276 AsnGluArgAsnSerLysLeuArgGlnGluAsnMetGluLeuAlaAlaAlaAlaAlaAla 295
 Db 819 GACATCCACATGCCCAACTGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 878
 QY 296 LeuLeuGluGlnTyrcysValGlnGluGluHisLleAspLysValPheLysHisLysAsp 315
 Db 879 CTTATTGACAGTATGCATTAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 938
 QY 316 LeuGlnGlnGlnLeuValAspAlaLysLeuGlnAlaGlnGluMetLeuLysGluAla 335
 Db 939 TTGCAACACAGCTTGTGGATGCCAACTTCAGCAACACACACAGCTGATAAAGAGCT 998


```

/db_xref="RZPD:DKFZp451A175"
/db_xref="taxon:9606"
/chromosome="6"
/clone="DKFZp451A175"
/tissue_type="human skeletal muscle"
/clone_lib="451 (synonym: hlccl). Vector pSport1; host DH10B; sites NotI + SalI"
/dev_stage="adult"
1..4154
/gene="DKFZp451A175"
<1..1518
/gene="DKFZp451A175"
/notes="similar to chicken 77K muscle-derived protein"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAD38617.1"
/db_xref="GI:21732876"
/db_xref="SPTREMBL:O8N3S2"
/translation="OKKLKLOKVOIQKEKDOLOGEHSRAILASRLSCLRELOR
HNKTLKEALQRAEEKEKEIYSHFQSLTDIQOIEQOSERNMELQENTLEAK
LSIIIDYELREBEHLDKIFHREIQKLVDAKLEQOMKGAERHREKEKELNQA
AEWKLQMLKEQDVLQAOITLYSGRFEFQSTLTKSNEVFATFKQEMDKTKKQQA
LEKDTAKWAKFENCOKALMDIMEKALRAKEYECFVMKIGLENCRALQERNELH
KKIRDAISIKDQNSDEEPESNVSDQIEAEVNSVQTAVKNLATAFMIHHP
ESTPHQSKETOPETIGSSQESADALKEPEOPPLIPSRDSSPLPLTPQAEAGGSDA
EPPSKASNPAGLGAETQCGLPVGAQDQPSMKPEAEASGOAPQAPTEASLQKMEAD
VPAPACAAAEHVAANVFACBPSRPPRAAAEELFVGNASQFPQPRNVADTNLEGVD"
3957..3962
/gene="DKFZp451A175"
3979
/gene="DKFZp451A175"

polyA_signal
polyA_site

ORIGIN
Alignment Scores:
Pred. No.: 4.89e-41 Length: 4154
Score: 1008.00 Matches: 203
Percent Similarity: 67.09% Conservative: 64
Best Local Similarity: 51.01% Mismatches: 71
Query Match: 37.31% Indels: 60
DB: 9 Gaps: 2

US-10-023-523-8 (1-530) x HSM803629 (1-4154)
QY 191 GlnLysGlnMetLysLeuLysGlnLysGlnSerGlnLysValGlnLysAspHis 210
DB 1 CAAAGAGATTAAAGCTCTCCAAAGAAACAGGTACAAATTCACAAAGAAAGAACCCAG 60
QY 211 LeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLysCysArg 230
DB 61 TTCAAGGTGAACACAGCAGAGCTATCTCTCGAAGCAATTTGGAGAGCTGTGTGCGCG 120
QY 231 GlnLeuGlnArgHisAsnArgSerLysGluGlyValGlnArgAlaArgGluGlu 250
DB 121 GAGCTGCAGACACACACAGACTCTGNAGAGAGAGCGCTTCAGCGGCGACGTGAGAA 180
QY 251 GlnGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeu 270
DB 181 GAAGAGAAAGAGAAAGAAATCACAAGCCATTTCCAGAGTAGTACCTTCACGGACATCCAGGGC 240
QY 271 GlnMetGlnGlnHisAsnGluArgAsnSerLysLeuArgGlnGlnAsnMetGluLeuAla 290
DB 241 CAGATCGAGCAGCAGCAGAGAGTATGAGGATATGAGCTCTGTCTGAGGAGACACAGAGCTTGA 300
QY 291 GlnArgLeuLysLysLeuIleGlnTyrGlnLeuArgGluGluHisIleAspLysVal 310
DB 301 GAAAGAGCTGAAAGCATCATCGATCAGTATGAGTTCAGAGAGGAGCATCTGGACAAATA 360
QY 311 PheLysHisLysAspLeuGlnGlnLeuValAspAlaLysLeuGlnAlaGlnGlu 330
DB 361 TTTAAACACAGAGAACTCAGCAGAGAGCTGTGTGATGCAAGAGCTTGAGCGGCCCAAGAA 420
QY 331 MetLeuLysGluAlaGluGluArgHisGlnArgGlnLysAspPheLeuLeuLysGluAla 350

```

```

DB 421 ATGATGAAGGAGCGGAGGAGGACACAAACGAGAAAGGAATATTTGCTGAACCGCA 480
QY 351 ValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlnThrHisLeuLysGlnGln 370
DB 481 CGAGAGTGGAAACTTCAGGCGGAAATGCTGAAGGACGAAGACAGAGTCTCTCAGGCTCAG 540
QY 371 LeuAlaLeuTyrThrGluLysPheGluPheGlnAsnThrLeuSerLysSerSerGlu 390
DB 541 CTCACCTCTACTCAGGAAGGTTTGAAGATTCAGAGACACTAATCTAAAGAAACGAG 600
QY 391 ValPheThrThrPheLysGlnGlnMetGluLysMetThrLysLysIleLysLysLeuGlu 410
DB 601 GTGTTTGCCAGCTTCAAAACAGGAAATGCAAAACAACTAAGAAATGAAGAGCTGAA 660
QY 411 LysGluThrThrMetTyrArgSerArgTrpGluSerSerAsnLysAlaLeuLeuGluMet 430
DB 661 AAGGACACAGCCACATGGAAGCCGATTTGAGAACTGTAAAGAGCTCTCTTTGACATG 720
QY 431 AlaGluGluLysThrValArgAspLysGluLeuGluGlyLeuGlnValLysIleGlnArg 450
DB 721 ATTGAAGAGAGCACTGAGACTTAAGATATGAGTCTTTGTGATAAATCGGAGG 780
QY 451 LeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgValGln 470
DB 781 CTAGAGAACTCTGCGGTCTTTACAGAGAGAGAAACGAACTCCACAAATAATCAGA 840
QY 471 AspLeuSerAlaGlyGlyGlnGlySerLeuThrAspSerGlyProGluArgArgProGlu 490
DB 841 GAGCAGAAATATCTGAAGAGATGACCAAGTCAGCAACTCCGATGAGAGCCAGAG 900
QY 491 Gly----- 491
DB 901 TCAAAAGCTCTCTGTGGATCAAGAGATTGACGAGAGAGGTTAATAGTGTCCAAACCGCC 960
QY 491 ----- 491
DB 961 GTGAAATATCTGCCACAGCCTTCATGATTAATTCATCATCAGAGTCAACCCGACCCAG 1020
QY 492 -----ProGlyAlaGln 495
DB 1021 TCCAAAGAAACCCAAACCGAATAGGAGTTCTCAGAGAGTGTGACGCGCTCTCAA 1080
QY 496 AlaProSerSerPro-----ArgValThrGluAlaProCysTyrProGly 510
DB 1081 GAGCCAGAGCAACCCCTCTGATCCCTTCACGGGATTCAGAGAGTCCCTGCTCCCTA 1140
QY 511 AlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAla 528
DB 1141 ACTCTCAGGCTGAAGCCAGGAGGAGGAGTGTGTAACCTCCCTCCAAGGCC 1194

RESULT 38
AX0533282 823 bp DNA linear PAT 13-JAN-2001
Sequence 48 from Patent WO0073801.
AX0533282
AX0533282.1 GI:12227601
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Obata, Y.
Breast, gastric and prostate cancer associated antigens and uses
therefor
Patent: WO 0073801-A 48 07-DEC-2000;
LUDWIG INSTITUTE FOR CANCER RESEARCH (US)
Location/Qualifiers
1..823
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

FEATURES
source
ORIGIN

```


Alignment Scores:
 Pred. No.: 1,046-37 Length: 823
 Score: 925.50 Matches: 197
 Percent Similarity: 86.55% Conservatives: 41
 Best Local Similarity: 71.64% Mismatches: 37
 Query Match: 34.25% Indels: 5
 DB: 6 Gaps: 0

US-10-023-523-8 (1-530) x AX053282 (1-823)

QY 163 LeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysTyrAlaGlu 182
 DB 2 CTAACACCCCTTTCAACCCAGAGGAGAGCTGGCAGCTCTCTGTAAGAAATATCTCTGAT 61

QY 183 LeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLysLysGlnSer 202
 DB 62 CTTCTCGAGAGAGAGGAGTGTTCAGAGCAATAGAGATCTTCGAGAGAGCAAGCC 121

QY 203 GlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArg 222
 DB 122 CAGATTGTGAAGAGAAAGTTCACTGCAGAGTGAACATAGCAAGCTATCTTGGCAAGA 181

QY 223 SerLysLeuGluSerLysLeuGluGluGlnArgHisAsnArgSerLysLysGluGlu 242
 DB 182 AGCAAGCTAGAACTCTTTTCAGAGAACTTCAGCGTCACAAATAGACGTTAAAGAGGAA 241

QY 243 GlyValGlnAlaArgGluGluGluLysArgLysGluValThrSerHisPheGln 262
 DB 242 AATATCGACAGCAGCAGAGAGGAGAGAACGAGCTAAAGAGCAACTGCACATTTCCAG 301

QY 263 ValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeu 282
 DB 302 ATTACCTTAATGAAATTCAGCCAGCTGGAGCAGCAGTCATCCACCAACCCAAACTC 361

QY 283 ArgGlnGluAsnMetGluLeuAlaGluArgLysLysLeuIleGluGlnTyrGluLeu 302
 DB 362 CGACAGAAACATTAAGCTGGGGAGAGCTAAAGAGCTCATCGAAGTACGCACTG 421

QY 303 ArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeuValAsp 322
 DB 422 AGGGAAGAGCACATTGATAGGTGTTCAACATAGAACTGCAACAGCAGCTCGTGGAT 481

QY 323 AlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluLysGlnArgGlu 342
 DB 482 GCCAAATCGAGCAACAGCACAACTGATAAAGAGCTGATGAAGAAATCGAGAGAG 541

QY 343 LysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGln 362
 DB 542 AGAGAGTTTATTAAGAGAGCGAGCAATCGAGGCACAAATACCAAAATGAAACAG 601

QY 363 GlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluPheGln 382
 DB 602 CAAGAAGTACACTA-AAACAGCAGCTTCTCTTTATATGATAGTTGAAGAAATTCAG 660

QY 383 AsnThrLeuSerLysSerGluValPheThrThrPheLysGlnGluMetGluLysMet 402
 DB 661 ACTACCATGCAAAACATGAACTGCTTCA-ACCTTCAGAGCAAGAAATGGGAAAGAT 719

QY 403 ThrLysLysLysLysLysLeuGluLysGluThrThrMetTyrArgSerArgTrpGluSer 422
 DB 720 GCCAAGAAATTA-AAAATGGGAAAGAAACCAATTAATTTGGGTTCCTCAATGGGAAAC 778

QY 423 SerAsnLysAlaLeuGluMetAlaGluGluLysThrValArg 437
 DB 779 AATATTAAACN-CTTTTGCA-ATGGNTGAAGAAAGAAACAGTCCGG 821

RESULT 39
 AX834462
 LOCUS 2194 bp DNA linear PAT 15-DEC-2003
 DEFINITION Sequence 1586 from Patent EP1347046.
 AX834462
 ACCESSION AX834462.1 GI:39920597
 VERSION

KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R., Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahara, K. and Masuho, Y.

TITLE Full-length cDNA sequences
 JOURNAL Patents: EP 1347046-A 1586 24-SEP-2003;
 Research Association for Biotechnology (JP)

FEATURES source
 1. 2194
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:
 Pred. No.: 4,366-32 Length: 2194
 Score: 820.50 Matches: 241
 Percent Similarity: 45.89% Conservatives: 38
 Best Local Similarity: 39.64% Mismatches: 112
 Query Match: 30.37% Indels: 217
 DB: 6 Gaps: 18

US-10-023-523-8 (1-530) x AX834462 (1-2194)

QY 1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 20
 DB 114 AAAGCAGCCAGCAGCAACCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 173

QY 21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40
 DB 174 GCGGCTCTGCAAGTAGAGCAGAGAGGTCCCGGAGAGAGAGAGAGAGAGAGAGAGAG 233

QY 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 60
 DB 234 GGGGCTCAAGCAGAGACCGCTCAGCTCTGGGCGCTTCGTGATGCTCTCTGAGAGTGA 293

QY 61 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGly 80
 DB 294 CGCCAACTGGAAGACATACCTGAGCACATCTGTGTGGACATAATACAGAGGGGGCCG 353

QY 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAla-----Glu 95
 DB 354 GAGGATGGGACACAGGGTGAGCGGCTGAACCCGAGATCTGCCCTCTAGTTTCATAGG 413

QY 96 -LysSerArgThrTyrValAlaArg-----AsnGlyGluProGluProThrProValVa 113
 DB 414 CCAAAGTGTATGATCGTGTGTGTCAGGACCTAGAGGGCGCTCCCTGACCCAGCTTCT 473

QY 113 L-----TyrGlyGluLys-----GluProSerLysGlyAs 123
 DB 474 TGCCATCTTCATCTCTGGGAACAAAGCTCTGTTTGGTTTGAGGGAGTGGTTTGG 533

QY 123 pProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHisArgArgPr 143
 DB 534 TTCCTTATCCCTCAGCGCTGAGACATAGAGGCTCTCTGGGCCCTACAGTGAGACACGA 593

QY 143 oGlnGlu-----LysLysLysAlaLysGlyLeuGlyLysGlu 155
 DB 594 TTCAAGAAATCTGAATACCCCGCTTTTCTCTCCCGCCCAAGGCAAAA-----AAGA 644

QY 155 uileThrLeuLeuMetGln-----ThrLeuAsnThrLe 166
 DB 645 CTTAGTACTACTCTGGAGAGAGAGAGTGCAGAGCTACAGGCCCTGCTGCTTGGATT 704

QY 166 uSerThrProGluGluLysLeuAlaLeuCysLysLysTyrAlaGluLeuLeuGluGlu 186
 DB 705 CAGCCCTCCCAAGACAGACAGACAGGACCGACCTCATCATCAACCAACTGGACTTAC 764

QY 186 uHis-----Argasn-----189
 DB 765 GCACCTCCCTCCCATCCMAAATAAGAGATTATTTCCCTTATTTACGAGTCCAG 824
 QY 190 -----SerGlnly 192
 DB 825 TTGATTATACCTTTGAAGTAGCACCTCAGTCCTTCACCTTCCTCCATCCCTCTCTC 884
 QY 192 sGlnMetLysLeuGlnLysLysGlnSerGlnLeuValGlnGlnLysAspHisLeu 212
 DB 885 ACCTGACACAGGCTCGAGCGCTCCTCTAGT-----AGCAGCAGACCATTCCTT 935
 QY 212 gGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArg-----230
 DB 936 GGGATGCACATGCTAGTCTTTCCTA-----GATATGGCAAGCTTTGCCCACTGAG 989
 QY 230 -----230
 DB 990 CTAGGCTGTATGTTCTTAGAGGCAATGTTTTTGCCCATTCCTCCCATTTACAGAGAAT 1049
 QY 231 -----GluLeuGlnArgHisAsnArgSerLeuLysGluGlyValGln-----245
 DB 1050 CAGGCACACAGAGTAGGGCTTCCACCCTAGGTGATCAATCCTGGGTCTAGATTT 1109
 QY 245 -----245
 DB 1110 GAGTGTGTTATGCTTCTTGGGAGCAGATTCATCCATAAACCATGTGCTTACC 1169
 QY 246 ArgAlaArgGluGluGlnLysArgLysGluValThrSerHisPheGlnValPhe 265
 DB 1170 AAGTCTGCTACTCTGGAGAGAACACGCTGAGGTGGA-----AAGCTGACCTT 1220
 QY 265 uAsnAspIleGlnLeuGlnMetGluGlnHis-----AsnGluArgAs 279
 DB 1221 CCAGAGACTGGGGCCCATGTTGTGTGTACACATGGAGTCCATCATCATGATGAGA 1280
 QY 279 nSerLysLeuArgGlnGlnMetGluLeuAla-----GluArg 292
 DB 1281 TGGGGGCTCGGCAAGATGCCCTGGTGTGGTGTGGTGTGGCTACCTGAGAAAGGAGG 1340
 QY 292 gLeuLysLysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAspLysValPhe 312
 DB 1341 CTTGACAGCGGCTGCTCCACATCTTTGTG---CAGCATATCCAAAGCTTTCAA 1397
 QY 312 sHisLysAspLeuGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnMetLe 332
 DB 1398 ACACAGGACCTACACAGCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1457
 QY 332 uLysGluAlaGluGluArgHisGlnArgGluLysAspPhe-----345
 DB 1458 AAAGAGGCGAAGAGCGCGCCAGCGGAGAGAGGATTTGTGAGCTCAGGCCCCAGGG 1517
 QY 345 -----345
 DB 1518 TTGGGTGGGGTGGGAGGAGACAGGCTGGGCTCTGGCTCAGCTCATAGCCGGTTATAT 1577
 QY 345 -----345
 DB 1578 GGGAGAAGTCTGCCAGACACAGGACAGATTCCTTGATGATGATGATGATGATGATG 1637
 QY 345 -----345
 DB 1638 CTCAGTGGGTCTGCTGT 1697
 QY 346 -----LeuLeuLysGluAlaValGluSerGlnArgMetC 357
 DB 1698 GATCTGGGTCTGCTGT 1757
 QY 357 ySgLeuMetLysGlnGlnThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGlu 377
 DB 1758 GTGAGTGTATGAGCAGCAGAGACCCACCTGAGGACACAGCTTGCCCTATACACAGA 1817

QY 377 ySgLeuGluPheGlnAsnThrLeuSerLysSerGluValPheThrThrPheLysG 397
 DB 1818 AGTTTGGAGGTTTCCAGAACACACTTTCCAAAGAGCGGAGTATTCCACCATTCAGC 1877
 QY 397 InGluMetGluLysMetThr 403
 DB 1878 AGGAGATGAAAAGGTAAC 1897
 RESULT 40
 AK096939 2194 bp mRNA linear PRI 15-JUL-2002
 LOCUS Homo sapiens cDNA FLJ39620 fis, clone SMINT2001114.
 DEFINITION AK096939
 ACCESSION AK096939
 VERSION 1 GI:21756551
 KEYWORDS oligo capping; fis (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Kanehori, K., Ishibaashi, T., Chiba, Y., Fujimori, K., Hiraoka, S.,
 Tanai, H., Watanabe, S., Ishida, S., Ono, Y., Hotuta, T., Watanabe, M.,
 Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
 Yamamoto, S., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,
 Kimura, K., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,
 Watanabe, M., Takahashi-Fujii, A., Oshima, A., Sugiyama, A.,
 Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuho, Y., Nagai, K.
 and Isogai, T.
 TITLE NEDO human cDNA sequencing project
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2194)
 AUTHORS Isogai, T. and Yamamoto, J.
 TITLE Direct Submission
 JOURNAL Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7
 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
 COMMENT NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:
 Research Association for Biotechnology (RAB); cDNA library
 construction: Helix Research Institute (HRI) (supported by Japan
 Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; clone selection for full insert sequencing: HRI and
 RAB; annotation: HRI and RAB.
 FEATURES
 source
 Location/Qualifiers
 1..2194
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="SMINT2001114"
 /tissue_type="small intestine"
 /clone_lib="SMINT2"
 /note="cloning vector: pME18SFL3"
 ORIGIN
 Alignment Scores:
 Pred. No.: 4,36e-32 Length: 2194
 Score: 820.50 Matches: 241
 Percent Similarity: 45.89% Conservative: 38
 Best Local Similarity: 39.64% Mismatches: 112
 Query Match: 30.37% Indels: 217
 DB: 9 Gaps: 18
 US-10-023-523-8 (1-530) x AK096939 (1-2194)
 QY 1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 20
 DB 114 AAAAGAGCCCGAGCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 173
 QY 21 AlaAlaProAlaValGluAlaGluGlyProGlySerGlnAlaProArgLysProGlu 40
 DB 174 GCGGCTCTCTGAGTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 233

QY	41	GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer	60	Db	1281	TGGGGGGCTGGGCAAAAGTGCCTGGTCTGTGGGCTGTGGGGCTACCTTGAGAAAGGGAGCG	1340
Db	234	GGGGCTCAAGCCAGACGGCTAGTCTGGGGCCCTTCGTGATGCTCTGAGAGCTGAGC	293	QY	252	gLeuLysLysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAspLysValPheLy	312
QY	61	ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyProGly	80	Db	1341	CTTGACAGCCGACTGCTCCACCACATCTTTGTTG---CAGCATATCGCAAAAGCTTCAA	1397
Db	294	CGCCAACTGGAACACATCTAGACACATATGTGTGTGACAAATAACACAGGGGGCCCCGGC	353	QY	312	sHisLysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMetLe	332
QY	81	GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAla	95	Db	1398	ACACAAGGACCTACACAACAGCAGCTGCTGGATGCCAAGCTCCAGCAGGCCAGGAGATGCT	1457
Db	354	GAGATGGGGCACAGGTGAGCGGTGAACCCGAGAGATCGCCCTCTTAGGTTTCATAGG	413	QY	332	uLysGluAlaGluGluArgHisGlnArgGluLysAspPhe	345
QY	96	LysSerArgThrTyrValAlaArg-----AsnGlyGluProGluProThrProValVa	113	Db	1458	AAAGGAGCAGAGAAGAGCGGCACCGCGGAGAGGATTTTGTGAGGCTCAGGCCCCAGGG	1517
Db	414	CCAAAGTGATGATCGTGTGTGACGACCTAGAGGGGCTCCCTGACCCACCCCTTTCT	473	QY	345	-----	345
QY	113	1-----TyrGlyGluLys-----GluProSerLysGlyAs	123	Db	1518	TTGGGGTGGGGTGGGAGGAGACAGCTGGGCTCTGGCTCAGCTCATAGCCGGGTTATAT	1577
Db	474	TGCCATATCTCATCTCTCGGAACAAAGCTGCTGTTGTTGAGGGGAGTGGTTGG	533	QY	345	-----	345
QY	123	pProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHisArgArgPr	143	Db	1578	GGGAGAAGTCTGGCCAGACCCAGGCACAGATTCTTTGAGTACCAAGTCTGAGAGCAAGC	1637
Db	534	TTCTTATCCCTCAGCGCTGAGACATAGAGGCTTCCTGGGCCACTACAGTGAGACACGAAC	593	QY	345	-----	345
QY	143	oGlnGlu-----LysLysLysAlaLysGlyLeuGlyLysG	155	Db	1638	CTCAGTGGGTCTGGTGTGTTGGCTTAAACCAACATAGCCCCCTGGGGGCTTCTGCAG	1697
Db	594	TTCAAGAATCTGAATACCCCGGTTTCTCTCCCGCCCAAGGCAAAA-----AAGGA	644	QY	346	-----LeuLeuLysGluAlaValGluSerGlnArgMetC	357
QY	155	uIleThrLeuLeuMetGln-----ThrLeuAsnThrLe	166	Db	1698	GATCTGGGTTCTGCTTGGAAATAGCTCTCTGAAAGAGCAGTAGAGTCCAGAGAGATGT	1757
Db	645	CTTAGTACTACTGTGAGAGAGGAGGTGCAGGACTACCAGCCCTGCTGCTTTCATTTA	704	QY	357	ysGluLeuMetLysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluL	377
QY	166	uSerThrProGluGluLysLeuAlaLeuCysLysLysTyrAlaGluLeuLeuGluL	186	Db	1758	GTGAGTGTATGAGCAGCAAGAGAGCCACCTGAAACCAAGCTTCCCTTATACACAGAGA	1817
Db	705	CAGCCCTCCCAGACAGACAGCAGCCCTCATCATACCCCAACTGGACTTACCTGTAG	764	QY	377	ysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysG	397
QY	186	uHis-----ArgAsn-----	199	Db	1818	AGTTTGAGGAGTTCCAGAACACACTTTCACAAAGCAGCGAGGTATTCCACCATTCAGC	1877
Db	765	GCACCTTCCCTTCCCATCCAAAAAATGGAGTATTTCCTTATTTTCAGCAAGTCCAG	824	QY	397	lnGluMetGluLysMetThr	403
QY	190	-----SerGlnLy	192	Db	1878	AGGAGATGAAAAGGTAAT	1897
Db	825	TTGATTTTACCTTTGAAGTAGCACCTGAGTCTTCACTTCTCTCCATCCTTCTCTC	884	RESULT 41			
QY	192	sGlnMetLysLeuLeuGlnLysGlnSerGlnLeuValGlnGluLysAspHisLeuAr	212	LOCUS	AL671759/c	141790 bp	DNA linear ROD 12-JUL-2002
Db	885	ACCTGACACAGCTCTGACGGCTCTCTAGT-----AGCAGGACAGCCATTCTCT	935	DEFINITION	Mouse DNA sequence from clone RP23-151A15 on chromosome 4, complete		
QY	212	gGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArg-----	230	ACCESSION	AL671759		
Db	936	GGGATGCACATGCTAGTCTTTCCTA-----GATATGGCAAGTCTTTCGCAACTGAG	989	VERSION	AL671759.10	GI:22138725	
QY	230	-----	230	KEYWORDS	HTG.		
Db	990	CTAGGCTGTTATGTTCTTAGAGGCAATGTTTTCGCCATTTCTCCCATTTACAGAGAT	1049	SOURCE	Mus musculus (house mouse)		
QY	231	-----GluLeuGlnArgHisAsnArgSerLeuLysGluGluValGln-----	245	ORGANISM	Mus musculus (house mouse)		
Db	1050	CAGGACACAGAGTAGGGCTTCAGCCCATAGGTGATCAATCTGGGTGACAGATTT	1109	REFERENCE	1 (bases 1 to 141790)		
QY	245	-----	245	AUTHORS	Corby, N.		
Db	1110	GAGTGTGTTTATGCTTGGCTTCTGGGAGCAGATTCATCCATAAACCATGTCTTACC	1169	TITLE	Direct Submission		
QY	246	ArgAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnValThrLe	265	JOURNAL	Submitted (12-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,		
Db	1170	AAGGTCTGACTCACTGGGAGAGAACGACGTGAGTTGGA-----AAGCTGACCTT	1220	COMMENT	Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk		
QY	265	uAsnAspIleGlnLeuGlnMetGluGlnHis-----AsnGluArgAs	279		humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk		
Db	1221	CCAGAGACTTGGGGCCCATGTTGTGTGTGACATGGAGTCCATCATCATGATTCAGA	1280		On Aug 8, 2002 this sequence version replaced gi:21732087.		
QY	279	nSerLysLeuArgGlnGlnAsnMetGluLeuAla-----GluAr	292		----- Genome Center		
					Center: Wellcome Trust Sanger Institute		
					Web site: http://www.sanger.ac.uk		
					Contact: humquerry@sanger.ac.uk		

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; information on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/c_elegans/wormpep RP23-151A15 is from the RPCT-23 Mouse PAC Library
 constructed by the group of Pieter de Jong.
 For further details see <http://www.chori.org/bacpac/home.htm>
 VECTOR: pBAC3.6.

FEATURES

source

Location/Qualifiers
 1. 141790
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="4"
 /clone="RP23-151A15"
 /clone_lib="RPCI-23"

ORIGIN

Alignment Scores:
 Pred. No.: 9.58e-28 Length: 141790
 Score: 771.00 Matches: 275
 Percent Similarity: 24.49% Conservative: 11
 Best Local Similarity: 23.54% Mismatches: 25
 Query Match: 28.53% Indels: 860
 DB: 10 Gaps: 9

US-10-023-523-8 (1-530) x AL671759 (1-141790)

QY 217 LysalaValLeuAaArgSerLys-----LeuGluSerLeuCysArgGluLeu 232
 DB 62348 AAATCGCTGTGGCAAGACACTTGGAACTAGTTAGGGGAGCTTTTGTCAAAGCTG 62289
 QY 233 GlnArgHisAsnArgSer---LeuLysGluGluGlyValGlnArgAlaArgGluGlu 251
 DB 62288 GCAACTATGGCAATCTTTCTTCAGGAGAGAGCGCTGAGCGAGCCCTGAGGAGAG 62229
 QY 252 GluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGln 271
 DB 62228 GAGAAGCGCAAGAAAGTGACTTCACACTTCAGGTGACACTGAATGACATTCAGTGCAG 62169
 QY 272 MetGluGlnHisAsnGluAtrcAsnSerLysLeuArgGlnGluAsnMetGluLeuAlaGlu 291
 DB 62168 ATGGACAGCATACAGAGGAACTCCAGCTGCCCGCAGAGATATGGAGTAGCCGAG 62109
 QY 292 ArgLeuLysLysLeuIleGluGlnTyrGluLeuArgGluGlu----- 305
 DB 62108 AGGCTCAAGAAGTTGATCGAGCAATACGAGCTTCGTCGAGGAGGTGAGCTCGAGGCGATTG 62049
 QY 305 ----- 305
 DB 62048 AGGCCAGAGCTTGGGAGTGTGTAGTGAGTGCCAGGGCCCTGGGGCCGCAACTCGAGCGCT 61989
 QY 305 ----- 305
 DB 61988 GGGCTTCTCTCTTCAAGAGATGTATCCTGTAGATAGAGAGCTGAATAAATGGAAG 61929
 QY 305 ----- 305
 DB 61928 AAGAGATTACACACACACACACAGTGTATTCCTGTGTGTGTGAACGGGACATCTGT 61869
 QY 305 ----- 305
 DB 61868 AGAGTATTAGTGTAGAGTGAGGCGTGTCTCGTCCCGAGCTCCGCATCAGACCTGTAC 61809

QY 305 ----- 305
 DB 61808 ATCAGACCATGCCANTCCCAAGACATGGGTGGGAGAGACCGGATGAAGTGCCTTGT 61749
 QY 305 ----- 305
 DB 61748 TCCTAGAGTTGGGGCTTTCTTGAAGGGAACCTCCTGACTCCATAAATTACCTTTTCTGCA 61689
 QY 306 -HisIleAspLysValPheLysHisLysAspLeuGlnGlnLeuValAspAlaLysLe 325
 DB 61688 GCATATCGCAAAAGCTTCTCAACATAGAGACCTCGCAGCAGAGCTAGTGGACGCCAAGCT 61629
 QY 325 uGlnGlnAlaGlnGluMetLeuLysGluAlaGluGluArgHisGlnArgGluLysAspPh 345
 DB 61628 CCAGCAGGCCCCAGGAGATGCTGAAGGAGGAGGAGCGGCACCGACGAGAGAGGAGT 61569
 QY 345 eLeu----- 346
 DB 61568 TGT-CAGGCTCAGGCCCCAGGCGCTGTGTGGGGCTGGAAGGAGTGGCTGGGTCTTGGC 61510
 QY 346 ----- 346
 DB 61509 TTAGTTCAATCTATTTCATTATGATGGGAGAGATTTTGCCTGCCCAATCCAGTAATTGCC 61450
 QY 346 ----- 346
 DB 61449 TGTCCAGACAGCAGTCTGAGAGCAGGTAGCCTCAGTGGGTCCAGTGTGTGGCTAAAGA 61390
 QY 346 ----- 346
 DB 61389 AGGTCAGACATACCTCAGGGGCTTCTGACAGGTTCTGGGTCTCTTGTGGAATAAGCT 61330
 QY 347 -LeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHi 366
 DB 61329 CCTGAAGGAAGCGGTGGAGTCCAGAGGATGTGGAGCTGTGAAGCAGCAGGAGAGCCCA 61270
 QY 366 sLeuLysGlnGlnLeu----- 371
 DB 61269 CCTAAGCAGCAGGTT-GAGACGGGAACTGACCTGAGTTTCCAGGATGTTTCCAAAGG 61211
 QY 371 ----- 371
 DB 61210 GCCCATCTCAGGGCAGTGGAGACAGGACTGTCTATAGCTGGCTTTGTGCCAGCTGGCT 61151
 QY 371 ----- 371
 DB 61150 GTGTGTGTGGCCCAAGTCTTTTATTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 61091
 QY 371 ----- 371
 DB 61090 CTGTGTAGCCCTGGCTGTCTTGAACCTTGTGTGTACACAGGCTGGCTGGACACAGA 61031
 QY 371 ----- 371
 DB 61030 GATTTCCTGTCTTAGCCTCCTAAGTAATGGGATTAATGTAGGCATGTGCCATCATGCC 60971
 QY 371 ----- 371
 DB 60970 AGGCATTTTGTGTCTTTTGAAGAACAGGTCTCTGAGTTTGTAGCTGGGTGACAACTAAC 60911
 QY 371 ----- 371
 DB 60910 CCAGAACTTACCTTGTAGACCCAGGCTCTCTGTGGCGCTTGAACCTGTGTGACAGTCTCTGA 60851
 QY 371 ----- 371
 DB 60850 CTCCAGATGATGAGGGCTTAGTACAGGTCTAGCTCAATCTGTCTACTTTTATTTT 60791
 QY 371 ----- 371
 DB 60790 GCTGTATTGATATGATATGAGTGTGTTCATGATGATATTTTCCCGACCTGTGTTC 60731
 QY 371 ----- 371

Db 16072 CCAGTGGTAGTCTCCAGTGGAGTCAAGGAGCAATGGRAGAAGGAATAAAGACCA 16131
QY 433 ----- 433
Db 16132 GAAAAAACGGTCCAGTGTATGCTGCTTTTACATGTAAAGCAGCCAGGTAGTTGT 16191
QY 433 ----- 433
Db 16192 GATTTCACAGCTTGTAAATAGAGAAGAAAGAACTAACGATGGAGCAGCACTGCAAGCCA 16251
QY 433 ----- 433
Db 16252 GACCTTGCTGAAAAGTTTTTGGGTTTTTTTGTCTTTTGTCTGCTGAATGTTTTTAGGTA 16311
QY 433 ----- 433
Db 16312 CGTTGTTCAATGAACCTCTCTGAGCTCTGAGGATGGTATTAGTAGTCTCTTTATAG 16371
QY 433 ----- 433
Db 16372 ATGAGACAGGCTCAAAAGTCAAGTCTTTGCCAAGTCAAGTGGTAGATAAATGGAGAA 16431
QY 433 ----- 433
Db 16432 TAGCTTATCTCAAGCGTGCCCTTTTCTGCACCATGTGCCCCACCTGACAGCCTAGT 16491
QY 433 ----- 433
Db 16492 CATGGCTTCAACTAGGACTGTTTCTTAAAGGGGGCCAGCTTTGGACTCGGTCTGCTCTCA 16551
QY 433 ----- 433
Db 16552 GCCTTGTAAAGTGTTCGCCCAAGTGTGTATGTTAGTGGAGGTTGATGGGCAAGG 16611
QY 434 ----- LysThrValArgAspLysGluLeuGluGlyLeu 444
Db 16612 CACTGAAGGTCTCATTTCTTCCTAGAAAACAGTCCGGGATAAGAACTGGAGGCGCTG 16671
QY 445 GlnValIysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAsp 464
Db 16672 CAGTAAATTCACCGCTGGAGAGCTGTGCGGGCACTGCAGACAGCGCAATGAC 16731
QY 465 LeuAsnIArgValGlnAspLeuSerAlaGlyGlnGlySerLeuThrAspSerGly 484
Db 16732 CTGAACAAGAGGTACAGGACCTGAGTGTGTGGCCAGGCTCCCTCACTACAGTGGC 16791
QY 485 ProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGlu 504
Db 16792 CTTGAGAGAGGCGCAGAGGGGCTGGGGCTCAAGCACCCAGCTCCCGGAGGTACAGAA 16851
QY 505 AlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGlu 524
Db 16852 GCGCTTGTCTACCCAGGAGCAGCAGCAGCAGAGAGCATCAGGCGAGACTGGGCTCAAGAG 16911
QY 525 ProThrSerIlaArgala 530
Db 16912 CCCACCTCCCGCAGGGCC 16929

RESULT 43

AX239608
LOCUS AX239608 22255 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 51 from Patent WO0164874.
ACCESSION AX239608
VERSION AX239608.1 GI:15797283
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.
AUTHORS Low density liprotein binding proteins and their use in diagnosing
TITLE and treating atherosclerosis

JOURNAL Patent: WO 0164874-A 51 07-SEP-2001;
Boston Heart Foundation, Inc. (US)
FEATURES Location/Qualifiers
source i. 22255
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 3,77e-26 Length: 22255
Score: 721.50 Matches: 224
Percent Similarity: 23.78% Conservative: 1
Best Local Similarity: 23.68% Mismatches: 1
Query Match: 26.70% Indels: 721
DB: 6 Gaps: 4

US-10-023-523-8 (1-530) x AX239608 (1-22255)

QY 305 GluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeuValAspAlaLys 324
Db 14093 CAGCATATCGACAAAGTCTTCAACACACAGGACCTACACAGCAGCTGGTGGATGCCAAG 14152
QY 325 LeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGlnArgGluLysAsp 344
Db 14153 CTCCAGCAGGCCAGGAGATGCTAAAGGAGCAGAGAGGGCCACCGCGGAGAGGAT 14212
QY 345 Phe ----- 345
Db 14213 TTTGTAGGCTCAGGCCCCAGGGTTGGGTGGGGGTGGGAGGAGACAGGCTGGGTCT 14272
QY 345 ----- 345
Db 14273 GGCTCAGCTCATAGCCGGTTATATGGGAGAAGTCTGGCCAGACCAGGCACAGATTCTT 14332
QY 345 ----- 345
Db 14333 GAGTACCAAGTCTGAGCAGAGGAAAGCCTCAGTGGGTCTGGTCTTGGCTAAACACCAA 14392
QY 346 ----- LeuLeuLys 348
Db 14393 CATAGCCCTCGGGGCTTCTCAGAGGATCTGGGTCTGTCTTGGAAATAGCTCTGAAA 14452
QY 349 GluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnThrHisLeuLys 368
Db 14453 GAGCAGTAGAGTCCAGAGGATGTGTAGCTGATGAAGCAGCAGAGAGACCCACCTGAAG 14512
QY 369 Gln ----- 369
Db 14513 CAACAGGTGAGAGCATATACTGACCTGTGCTTCAAGTTTCCCTCACTGGGCCCAT 14572
QY 369 ----- 369
Db 14573 CTTGGGGTGTGAAATGGGACCCTCATTTCTAGGACTGGTGTGTCTGGTGTGTATGAC 14632
QY 369 ----- 369
Db 14633 GCCTTGGTTGAGCTTAGTGGGCTCAGAGGACTTCATTGTAGCTCAGAAATGTATTGCT 14692
QY 369 ----- 369
Db 14693 TTTGAGGAGGTAGGAACAGAGAGTTTGAAATCAACATAAAGGCAAAATAAAGTCACC 14752
QY 369 ----- 369
Db 14753 CTAAGTCTCTACTTTCCAGGCTTAGCATTTTGGATTATATCTCTTCCAATATATAGCTT 14812
QY 369 ----- 369
Db 14813 TCGTTTGTTTTAAAGGAATAATAGTATCTCAATAGATTACTTGGTCAGAGAGTCAGGACG 14872
QY 369 ----- 369

ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 529)
AUTHORS Jacobs, K., McCoy, J.M., Lavallie, E.R., Racie, L.A., Merberg, D.,
Treach, M., Spaulding, V. and Agostino, M.J.
TITLE Secreted expressed sequence tags (seSTs)
JOURNAL Patent: JP 2001518793-A 869 16-OCT-2001;
GENETICS INSTITUTE INC
COMMENT PN JP 2001518793-A/869
PD 16-OCT-2001
PF 10-APR-1998 JP 1998543070
PR 10-APR-1997 US 08/837312
PI KENNETH JACOBS, JOHN M MCCOY, EDWARD R LAVALLIE, LISA A RACIE, PI
DAVID MERBERG,
PI MAURICE TREACH, VIKKI SPAULDING, MICHAEL J AGOSTINO PC
C12N15/12.C12N5/10.C07K14/47.C12Q1/68.A61K38/17 CC Strandedness:
Double;
CC Topology: Linear; Location/Qualifiers.
PH Key Location/Qualifiers.
FEATURES
source
i. 529
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"
ORIGIN
Alignment Scores: 1,94e-27 Length: 529
Pred. No.: 713.00 Matches: 137
Score:
Percent Similarity: 99.28% Conservative: 1
Best Local Similarity: 98.56% Mismatches: 0
Query Match: 26.39% Indels: 0
Gaps: 6
DB:
US-10-023-523-8 (1-530) x BD060509 (1-529)
QY 1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 20
DB 104 ARAAGCAGCCAGGACCAACCGGAAGCAGGACCCGAGGAGCCGAGGCGGCCAGCCAG 163
QY 21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40
DB 164 GCGGCTCTCTAGTAGAAGCAGGAGGTCCGCGCAGCAGCCAGCTCTCTGGAGCCGCGAG 223
QY 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 60
DB 224 GGGGCTCAAGCCAGACGAGCTCAGTCTGGGGCCCTTCGTGATCTCTGAGGAGCTGAGC 283
QY 61 ArgGlnLeuGluAspLeuSerThrTyrCysValAspAsnAsnGlnGlyProGly 80
DB 284 CGCCAACTGGAAGACATACCTAGCAGCATACTGTGTGACAAATACACAGGGGGCCCCGGC 343
QY 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
DB 344 GAGATGGGACACAGGTGAGCGGCTGACCGCTGAACCCGAGATCCAGAGTCCCGGACCTAT 403
QY 101 ValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGluLysGluProSer 120
DB 404 GTGGCAAGGAATGGGAGCGCTGAACCACTCCAGTAGTCAATGAGGAGAGGAACCTCC 463
QY 121 LysGlyAspProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAsp 139
DB 464 AAGGGGATCAAAACAGAGAGATCCGGCAGAGTCCGAGGTCGAGAGCCGAGAA 520
RESULT 45
LOCUS BC046565 3554 bp mRNA linear PRI 30-SEP-2003
DEFINITION Homo sapiens taxilin, mRNA (cdna clone IMAGE:5534975), complete cds.
ACCESSION BC046565
VERSION BC046565.1 GI:28302155

KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 3554)
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Udgin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullany, S.J., Bosak, S.A., McSwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
TITLE Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE 22388257
PUBMED 12477932
REFERENCE 2 (bases 1 to 3554)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (10-FEB-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@hgrl.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Green, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,
Madero, O.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Tsurgon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, D.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 78 Row: p Column: 11
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis
This clone has the following problem: The cds is short compared to
the longest cds in the locus.
Location/Qualifiers
1. 3554
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5534975"
/tissue_type="Uterus, leiomyosarcoma"
/clone_lib="NIH MGC 71"
/lab_host="DH10B"


```

Qy 261 eGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLy 281
Db 663 TCAGTGACTATACCGCGTGTGAGTCCAAATGAACACTACAGTTGAAGTGCACACAA 722
Qy 281 sLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLeuLeuLeuGlnIleGlnTyrl 301
Db 723 AATGAGTTCGAAATGATCAGATTAGTTCACAAATCAATGATTTGGTTGAAAGTTGAA 782
Qy 301 uLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnGlnLeuVa 321
Db 783 TCAAAGAGAGTGTGATTGAGATCTTGTGAAGAATCTCGCTCCAGAGCAACTTAG 842
Qy 321 lAspAlaLysLeuGlnGlnGlnGlnMetLeuLysGluAlaGluGluArgHisGlnAr 341
Db 843 CTCATCGAGTTAAAGAAAGATCACCATGGAGCACAGATGCAAAAGGAGGCAGATCGCA 902
Qy 341 gGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLy 361
Db 903 TGAGAAAGTCGTCTCTTCTGGAACGAATAAAGAGATGAATGAATCTTATGAACGATGAT 962
Qy 361 sGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyThrGluLysPheGluGluPh 381
Db 963 CAACACGAAAGGATTGAAATTCAGCTCCAGTATTATCTGAAAAATTTGAGGACTT 1022
Qy 381 eGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLy 401
Db 1023 TAAGTCTGTTTGAAGAAAGCAATGATATTTTCACTTCCTATCAAACTGAGCTGAAAA 1082
Qy 401 sMetThrLysLysIleLysLysLeuGluLysGluThrThrMetTyArgSerArgTrpGl 421
Db 1083 TCTTGGAAACAGTTAAAGAGGAGAGAGAAAGAAACACACATACAGAAAAATGGGA 1142
Qy 421 uSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGluLe 441
Db 1143 GGATTCGACAAAGCACTTGTGATCTGTGAA-----TATAC 1181
Qy 441 uGluGluLeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGl 461
Db 1182 CCAGCAGCTTGAACGAAACAGCAGCTGTGTGAAAGTTAAGTCGAGCCTTCAGACCGA 1241
Qy 461 uArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuth 481
Db 1242 ACCTAATGAGTTGTACAGAGGTTGCAAGATGATGA-----CGGAAAC 1286
Qy 481 rAspSerGlyProGluArgArgProGluGluProGlyAlaGlnAlaProSerSerProAr 501
Db 1287 ATATGAAGAAACCAAGAAACT-----GCGGATGCTCTGTCTTCCAG 1334
Qy 501 q 501
Db 1335 A 1335

RESULT 47
HSDJ622L5 115756 bp DNA linear PRI 22-NOV-2001
LOCUS Human DNA sequence from clone RP4-622L5 on chromosome 1p34.2-36.11,
DEFINITION complete sequence.
ACCESSION AL049795
VERSION AL049795.21 GI:17065925
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 115756)
Frankland, J.
Direct Submission
Submitted (22-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Nov 25, 2001 this sequence version replaced gi:6010175.
During sequence assembly data is compared from overlapping clones.

```

Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information

on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr1

This sequence is the entire insert of clone RP4-622L5 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP4-622L5 is from the library RPCI-4 constructed by the group of Pieter de Jong. For further details see

http://www.chori.org/bacpac/home.htm

VECTOR: pCYPAC2.

FEATURES

source

Location/Qualifiers

1..115756

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="1"

/map="p34.2-36.11"

/clone="RP4-622L5"

/clone_lib="RPCI-4"

1645..1810

/note="TIGER1 repeat: matches 2259..2417 of consensus"

2123..2437

/note="TIGER1 repeat: matches 1957..2259 of consensus"

2726..2782

/note="TIGER1 repeat: matches 1904..1957 of consensus"

3067..3486

/note="TIGER1 repeat: matches 1501..1904 of consensus"

3781..4180

/note="TIGER1 repeat: matches 1117..1501 of consensus"

5053..5082

/note="TIGER1 repeat: matches 1090..1117 of consensus"

5372..5501

/note="TIGER1 repeat: matches 966..1090 of consensus"

5628..5704

/note="TIGER1 repeat: matches 893..966 of consensus"

5991..6492

/note="TIGER1 repeat: matches 411..893 of consensus"

6800..6835

/note="TIGER1 repeat: matches 378..411 of consensus"

6968..7022

/note="TIGER1 repeat: matches 1..378 of consensus"

8281..8358

/note="3 copies 26 mer 79% conserved"

8301..8360

/note="30 copies 2 mer ta 81% conserved"

17674..17777

/note="4 copies 26 mer 77% conserved"

17685..17784

/note="50 copies 2 mer at 82% conserved"

19385..19542

/note="match: GSS: Em:AQ306911"

19551..19574

/note="12 copies 2 mer tt 100% conserved"

20865..20922

/note="29 copies 2 mer at 70% conserved"

20934..21043

/note="55 copies 2 mer at 70% conserved"

```

repeat_region 20940..21043
repeat_region /note="4 copies 26 mer 74% conserved"
misc_feature 21168..21234
misc_feature /note="tRNA-Asp-GAC repeat: matches 1..75 of consensus"
misc_feature 28263..28691
misc_feature /note="match: GSS: Em:AQ317905"
misc_feature complement(29964..30447)
misc_feature /note="match: GSS: Em:AQ732256"
misc_feature 30465..30718
misc_feature /note="match: STS: Em:G58909"
misc_feature match: GSS: Em:AQ376964"
repeat_region 38923..39446
repeat_region /note="L1MEC repeat: matches 1546..2095 of consensus"
misc_feature 39761..39882
misc_feature /note="L1MEC repeat: matches 1429..1546 of consensus"
repeat_region complement(43254..43726)
misc_feature 45555..46050
misc_feature /note="MER44C repeat: matches 1..728 of consensus"
repeat_region 46055..46468
misc_feature /note="match: GSS: Em:AQ040362"
repeat_region 48914..48982
misc_feature /note="MER21B repeat: matches 718..790 of consensus"
repeat_region 48985..49173
misc_feature /note="MER21B repeat: matches 77..263 of consensus"
misc_feature 49285..49651
misc_feature /note="match: GSS: Em:AQ761532"
misc_feature 52462..53177
misc_feature /note="match: GSS: Em:AQ040082"
misc_feature 54716..54865
misc_feature /note="match: GSS: Em:AQ680771"
repeat_region 60551..60614
misc_feature /note="32 copies 2 mer tt 70% conserved"
misc_feature complement(62219..62382)
misc_feature /note="match: STS: Em:G20585"
misc_feature complement(66030..66267)
misc_feature /note="match: GSS: Em:AQ514287"
misc_feature complement(66041..66267)
misc_feature /note="match: GSS: Em:AQ803288"
misc_feature 66066..66255
misc_feature /note="match: GSS: Em:B69239"
misc_feature 68882..69541
misc_feature /note="CpG island"
repeat_region /evidence=not_experimental
misc_feature 71427..71448
misc_feature /note="11 copies 2 mer tt 100% conserved"
misc_feature 73527..74249
misc_feature /note="match: GSS: Em:AQ892680"
misc_feature 74514..74926
misc_feature /note="match: GSS: Em:AQ441232"
misc_feature 74532..74981
misc_feature /note="match: GSS: Em:AQ812740"
misc_feature 74626..75008
misc_feature /note="match: GSS: Em:AQ091614"
misc_feature complement(74682..75020)
misc_feature /note="match: STS: Em:L29744"
misc_feature complement(74683..74951)
misc_feature /note="match: STS: Em:L29729 Em:L29763"
misc_feature 74717..74867
misc_feature /note="match: STS: Em:G15744"
misc_feature 82692..83089
misc_feature /note="match: STS: Em:G09447"
misc_feature 83391..83481
misc_feature /note="match: GSS: Em:A2310334"
repeat_region 85284..85337
misc_feature /note="27 copies 2 mer tt 88% conserved"
misc_feature 85339..85718
misc_feature /note="match: STS: Em:G28385"
misc_feature complement(87250..87612)
misc_feature /note="match: STS: Em:G06968"
misc_feature complement(87455..87579)
misc_feature /note="match: STS: Em:G26511"
misc_feature complement(87851..88313)

/note="match: GSS: Em:AQ536603"
89654..89685
/note="16 copies 2 mer tc 87% conserved"
91856..92386
/note="match: GSS: Em:B85187"
92101..92152
/note="26 copies 2 mer tt 71% conserved"
94297..95612
/note="CpG island"
/evidence=not_experimental
complement(98493..98894)
/note="match: GSS: Em:AQ207463"
103719..104155
/note="match: GSS: Em:AQ543924"
104406..104437
/note="MERS7B repeat: matches 303..334 of consensus"
107785..107841
/note="L1MB1 repeat: matches 6111..6171 of consensus"
111249..111767
/note="CpG island"
/evidence=not_experimental
113362..113450
/note="match: STS: Em:G16146"
match: GSS: Em:AQ806873"

ORIGIN
Alignment Scores:
Pred. No.: 6.57e-23 Length: 115756
Score: 671.00 Matches: 223
Percent Similarity: 23.26% Conservative: 1
Best Local Similarity: 23.16% Mismatches: 2
Query Match: 24.83% Indels: 739
DB: 9 Gaps: 4

US-10-023-523-8 (1-530) x HSDJ622L5 (1-115756)
QY 305 GluHisIleAspLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLys 324
Db 81636 CAGCATATCGACAAAGTCTTCAACACACAGGACCTTCAACAGCAGCTGGTGGATGCCAAG 81695
QY 325 LeuGlnGlnAlaGlnGlnMetLeuLysGluAlaGluGluArgHisGlnArgGluLysAsp 344
Db 81696 CTCACAGGCGCCAGGAGATGCTAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAT 81755
QY 345 Phe-----345
Db 81756 TTGTGAGGCTCAGGCCCCCAGGGTTGGGGTGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAT 81815
QY 345 -----345
Db 81816 CTCAGTCTATAGCCGGGTTATATGGGAGAGTCTGGCCAGACACAGGACAGATCCTTGA 81875
QY 345 -----345
Db 81876 GTACAGTCTGAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 81935
QY 346 -----346
Db 81936 TAGCCCTCGGGGCTTCTGACAGGATCTGGGGTCTGTCTTGGAAATAGCTCTCTGAAAGA 81995
QY 349 uAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLeuLysG1 369
Db 81996 GGCAGTAGAGTCCAGAGATGTGTAGCTGTATGAAGCAGCAGAGAGAGAGAGAGAGAGAGAGAG 82055
QY 369 nGlnLeu-----371
Db 82056 ACAGGT-GAGAGCATATACCTGACCTGTGCTTCAAGTTTCCCTCACTGGGCCCATC 82114
QY 371 -----371
Db 82115 CTGGGGTAGTGAATGGAGCCCTCATTTAGGACTGGCTGTGTCTCTGCTCTATGACG 82174
QY 371 -----371

```

Db	82175	CCTTGGTTAGCCCTTTGTTCTCTCCGGACCTGACAGTACCTATGATGGTGGTGACACAGGT	82234	Db	83254	TGACTGCATCCTGTAAATGCCTTGGCCTTGGGATCAATCATTTCCACCTTTGGAGACACA	83313
QY	371	-----	371	QY	401	-----	401
Db	82235	AGTTAGTGGGCTCAGAGGACTTCATTTGTAGCTCAGAAATGATTGCTTTTGAGGAGGT	82294	Db	83314	GCTGAGTCCCACTTGGAGACACAGACCTTGGAGAGCCAGCTCTGACCATTTCCCTTC	83373
QY	371	-----	371	QY	402	-----	415
Db	82295	AGGAACAGAGAGTTTGAATAACAAATAAGGCAAAATAAAGTCACCTCAAGTCTCT	82354	Db	83374	TGTCTGTACATAACCTAGATGACTAGAGATCAAGAGCTGGAGAAAGAAACCAT	83433
QY	371	-----	371	QY	415	-----	433
Db	82355	ACTTTCAGGCTTAGCATTTTGGATTATATCTTCCAAATATATACCTTTGCTTTGTTTT	82414	QY	433	-----	433
QY	371	-----	371	Db	83494	CTGTCTGTGATCTGCAGCCAGGCTGGGGTGTGCATCTAGCGCATATCAGGCCCTTTCCCT	83553
Db	82415	AAGGAAAAATAGTATCTCATAGATTAAGTGTGAGAGTCAAGACGGGTCTGAGTGT	82474	QY	433	-----	433
QY	371	-----	371	Db	83554	GTATGTTCTACCCATCAGTGACACAGCTAGCATGAGGTAGAGGTGAGATTTGCACACAAT	83613
Db	82475	GTTGACGAGTGCCTCCCGAGAAACCCAGTCTTATCTGTGGGCTGCTTTCTCCCCACA	82534	QY	433	-----	433
QY	372	-----	390	Db	83614	GTCCAGTCCAAAGTTAATGCTGTTCTCTCCCATGGAGGTGGAGCCCAAGTGGTAGG	83673
Db	82535	GCTTGCCCTATACACAGAGAAGTTTGGAGGTTTCCAGAACACACTTTCCAAAAGCAGCA	82594	QY	433	-----	433
QY	390	uValPheThrPheLysGlnMetGluLys-----	401	Db	83674	TCTCCAGTGGAGTGAAGGAGCAAAATGGAAGAAAGGAAATAAAGACAGCAGAAAAACGG	83733
Db	82595	GGTATTCAACACATTCAGCAGGAGATGGA-AAAGGTAACTGTGTCCAGGCCAGGCATG	82653	QY	433	-----	433
QY	401	-----	401	Db	83734	GTGCCAGTGATGTGCTGCTTTTACATGTAAAGCAGCCAGCTAGTTGTGATTTTACAGC	83793
Db	82654	GCTGCTGGGCATAAGCTGCTTCATTCAAAATTTGGGCTGCCCTCAGAGCTCCCA	82713	QY	433	-----	433
QY	401	-----	401	Db	83794	TTGTAATGTAGAAAGAACTAACTGATGGAGCAGCAACTGCAAGCCAGACCTTGTCTGA	83853
Db	82714	TCTGGGTGTCTCAAGGGCAGGCTGTAGGAAGTTTACAGCCTTTCCCTCTTGAGGC	82773	QY	433	-----	433
QY	401	-----	401	Db	83854	AAATTTTGGGTTTTTTTGTCTTTTGTCTGCTGAATGTTTTTAGGTACGTGTCTATT	83913
Db	82774	AGTATCAGTGGTATGATACACTCCAGGTTGTCCAGGGAATGGGCGAGTCTTTCTGTT	82833	QY	433	-----	433
QY	401	-----	401	Db	83914	GAACCTTCTCTGAGCTCTGAGGATGATTTAGTAGTCTCTGTTTTATAGATGAGACAGC	83973
Db	82834	TGTTTGGTTTTTTTGGGGTTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	82893	QY	433	-----	433
QY	401	-----	401	Db	83974	TCAAAAGTCAAGTCTCTTTGCCAAGTCACTGCTGATGATTAATGGAGGAATAGCTTATCTC	84033
Db	82894	GAGACTCACCTATTGCCAGGCTGGAGTGCAGTGGCATGATCTCAGCTCATTGACGCTT	82953	QY	433	-----	433
QY	401	-----	401	Db	84034	CAAGCGTGCCTTTTCTGCACCATGCTGCCACCTGACAGCCTAGTCTGCTTCA	84093
Db	82954	TGCCCCCGGGTTCAAGTGAATCTCTGCTCAGCTCAGCTCTGACTAGCTGGATTACAGC	83013	QY	433	-----	433
QY	401	-----	401	Db	84094	CTAGGACTGTTTCTTAGAGGGGGCCAGCTTTGGACTCGTCTGCTCAGCCTTGTAAA	84153
Db	83014	GCCTGCCACCATGCTGGCTAATTTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	83073	QY	433	-----	433
QY	401	-----	401	Db	84154	GTGTTTGGCCCAAGTGTGTGATGTTAGTGGGAGGTTGATGGGCGACGCACTGAAGTCT	84213
Db	83074	GGGGTTTCAACATGTTGGCCAGGCTGTCTCGAACTCTTGGCCTCAAGTGAATCTGCCCG	83133	QY	434	-----	448
QY	401	-----	401	Db	84214	TCATTTCTTCTCCTAGAAAACAGTCCGGGATAAAGAACTGGAGGCTCTGACAGTAAAT	84273
Db	83134	CTTGGCTCCAAAGTCTGGATTATAGCGGTGAGCCACCATGCTGGCCCTTACCAT	83193	QY	448	-----	468
QY	401	-----	401	Db	84274	CCACCGCTGGAGAGCTGTGCCGGGCACTGCAAGCAGAGCGCATGACCTGAACAGAG	84333
Db	83194	TCCTTGTATTGTTGGTGGACACCTCTGACTTCTCTGTTGGTGGTGGCAGAGGGCAT	83253	QY	468	-----	488
QY	401	-----	401	Db	84334	GGTACAGACCTTGAAGTGTGTTGGTGGCCAGGGTCTCCCTCACTCAGAGTGGCCCTGAGAGG	84393

```

QY 488 gProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTy 508
Db 84394 GCCAGAGGGGCTGGGGCTCAAGCACCCAGCTCCCCCAGGGTCAAGAGCGCTTGCTA 84453

QY 508 xProGlyValProSerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAl 528
Db 84454 CCCAGGAGCCCGAGCAGAGCATCAGGCCAGACTGGGCCCTCAAGAGCCCACTCCGC 84513

QY 528 aArgAla 530
Db 84514 CAGGGCC 84520

RESULT 48
BC029686 3495 bp mRNA linear PRI 20-MAY-2002
LOCUS BC029686 Homo sapiens, clone IMAGE:4993857, mRNA, partial cds.
DEFINITION BC029686
ACCESSION BC029686
VERSION BC029686.1 GI:20987361
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3495)
Strausberg, R.
Direct Submission
Submitted (06-MAY-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: csapsb-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amgebcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 42 Row: j Column: 8.
Location/Qualifiers
1. 3495
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4993857"
/tissue_type="placenta, choriocarcinoma"
/clone_lib="NIH MGC 10"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
<1..393
/codon_start=1
/product="unknown (protein for IMAGE:4993857)"
/protein_id="AAH29686.1"
/db_xref="GI:20987362"
/translaton="MTTKIKLEKTTMYRSWESNKKALEMAEKTVRDKELEGL
QVQIQLEKLCRALQTERNDLNKRVDLSAGGGGSLTDSGPERRRPGQAQPSRPV
TEAPCYGAPSTASGQTGPQPTESARA"

FEATURES
source
CDs
Alignment Scores:
Pred. No.: 2,93e-24 Length: 3495
Score: 566.00 Matches: 130
Percent Similarity: 100.00% Conservative: 0

ORIGIN

```

```

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 24.65% Indels: 0
DB: 9 Gaps: 0
US-10-023-523-8 (1-530) x BC029686 (1-3495)

QY 401 LysMetThrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 420
Db 1 AAGATGATCTAGAAGATCAAGAGCTGGAGAAAGAACCAACCATGTACCGGTCCGGTGG 60

QY 421 GluSerSerAsnLysAlaLeuLeuGluMetAlaGluLysLysThrValArgAspLysGlu 440
Db 61 GAGACGACCAACAGGCCCTGCTTGAGATGCTGAGGAGAAACAGTCCGGGATAAAGAA 120

QY 441 LeuGluGlyLeuGlnValLysLysLysLysLysLysLysLysLysLysLysLysLysLys 460
Db 121 CTGGAGGGCCCTGCAAGGTAAATAATCAACGGCTGAGAGAGCTGTCCCGGGCACTCGACACA 180

QY 461 GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeu 480
Db 181 GAGCCCATGACCTGAACAGAGGGGTACAGACCTGAGTGTGCTGGCCAGGGGCTCCCTC 240

QY 481 ThrAspSerGlyProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerPro 500
Db 241 ACTGACAGTGGCCCTGAGAGGAGCCAGAGGGGCTGGGGCTCAAGCACCCAGCTCCCCC 300

QY 501 ArgValThrGluAlaProCysTyProGlyAlaProSerThrGluAlaSerGlyGlnThr 520
Db 301 AGGCTCAGAGAGCCCTTGTACCCAGGACACCGAGCACAGAGCATCAGGCCAGACT 360

QY 521 GlyProGlnGluProThrSerAlaArgAla 530
Db 361 GGGCCTCAAGAGCCACCTCCGCCAGGGCC 390

RESULT 49
AY047502 1841 bp mRNA linear INV 16-AUG-2001
LOCUS Drosophila melanogaster GH01188 full length cDNA.
DEFINITION AY047502
ACCESSION AY047502
VERSION AY047502.1 GI:15010371
KEYWORDS FLI CDNA.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Peerygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1841)
Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Carlson, J.,
Champe, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George, R.,
Gonzalez, M., Guarin, H., Li, P., Liao, G., Miranda, A., Mungall, C.J.,
Nunoo, J., Pacleb, J., Paragas, V., Park, S., Phouanavong, S., Wan, K.,
Yu, C., Lewis, S.E., Rubin, G.M. and Celniker, S.
Direct Submission
Submitted (19-JUL-2001) Berkeley Drosophila Genome Project,
Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory
Berkeley, CA 94720
This clone was sequenced as part of a high-throughput process to
sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
Science 2000). The sequence has been subjected to integrity checks
for sequence accuracy, presence of a polyA tail and contiguity
within 100 kb in the genome. Thus we believe the sequence to
reflect accurately this particular cDNA clone. However, there are
artifacts associated with the generation of cDNA clones that may
have not been detected in our initial analyses such as internal
priming, priming from contaminating genomic DNA, retained introns
due to reverse transcription of unspliced precursor RNAs, and
reverse transcriptase errors that result in single base changes.
For further information about this sequence, including its location
and relationship to other sequences, please visit our Web site
(http://fruitfly.berkeley.edu) or send email to

```


Page 60

Page 60


```
Db 143119 ATATCCACCCTGCCTCTACATCGAGTGTGGGATAAAGTGGGGCTGGTGTGTATGGGTA 143178
Qy 401 ----- 401
Db 143179 TGCAGACGCCACAGATGCATATGGAGGCCCGAGGACAATTACAGGAGCGGACTCTGCCG 143238
Qy 401 ----- 401
Db 143239 GTTTGGGGCATGGGGCTCATTAGGCTGGCGAGGCATCAAGTGCCTTCTTTATTTCCTAAG 143298
Qy 401 ----- 401
Db 143299 CAGTCTTGTGCGACCCCTCAGTGCTCAGATTAACTCGAGGCTCCCGAGTTGGCACAGAGTAA 143358
Qy 402 -----MetThrLysLysIleL 407
Db 143359 AGCCAGTCCAACCCCTGACCTTCTGTTCATATGTGTAAACAGATGACAAAGAGATCA 143418
Qy 407 YsLysLeuGluLysGluThrThrMetTyrArgSerArgTTPGluSerSerAsnLysAlaL 427
Db 143419 AGAAGCTGGAGAAAGAGACACCCTCTACCCGGTCCCGGTGGGAAAGCAGCAACAGGCTC 143478
Qy 427 euLeuGluMetAlaGluGlu 433
Db 143479 TGCTGGAGATGGCTGAGGAG 143498
```

Search completed: June 8, 2004, 10:32:38
Job time : 6611.15 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 8, 2004, 06:37:58 ; Search time 113.29 Seconds

(without alignments)
2596.205 Million cell updates/sec

Title: US-10-023-523-8

Perfect score: 2702
Sequence: 1 KSPGQFPAAGGAGQRRPQ.....APSTASGQTGPBTSARA 530

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 50 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10023523/runat_08062004_063749_29963/app_query.fasta_1.1422
-DB=Issued Patents NA -OPM=fastap -SUFFIX=rni -MINMATCH=0.1 -GOOPCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=50 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=50
-MODE=LOCAL -OUTFWT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USR=US10023523 @CGN 1.1.98 @runat_08062004_063749_29963 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2702	100.0	4697	4	US-08-979-608A-17
2	2702	100.0	4697	4	Sequence 17, Appl
3	2702	100.0	4697	4	Sequence 17, Appl
4	2693	99.7	1638	4	Sequence 17, Appl
5	2465.5	91.2	4722	4	Sequence 46, Appl
6	2465.5	91.2	4722	4	Sequence 14, Appl
7	2465.5	91.2	4722	4	Sequence 14, Appl
8	1597	59.1	1854	5	Sequence 14, Appl
9	1047.5	38.8	1191	2	Sequence 1, Appl
10	721.5	26.7	22255	4	Sequence 5, Appl
11	652.5	24.1	614	4	Sequence 51, Appl
12	278.5	10.3	2680	2	Sequence 1333, Ap
					Sequence 5, Appl

13	278.5	10.3	2680	2	US-08-742-923A-5	Sequence 5, Appl
14	278.5	10.3	2887	2	US-08-533-306A-3	Sequence 3, Appl
15	278.5	10.3	2887	2	US-08-742-923A-3	Sequence 3, Appl
16	274	10.1	9551	1	US-08-056-200-93	Sequence 93, Appl
17	274	10.1	9551	2	US-08-800-644-93	Sequence 93, Appl
18	268.5	9.9	5661	4	US-09-023-655-548	Sequence 548, App
19	268.5	9.9	5661	4	US-08-938-105-2	Sequence 2, Appl
20	268	9.9	8948	4	US-09-643-597-119	Sequence 119, App
21	268	9.9	8948	4	US-09-480-884A-119	Sequence 119, App
22	268	9.9	8948	4	US-09-542-615A-119	Sequence 119, App
23	268	9.9	8948	4	US-09-606-421B-119	Sequence 119, App
24	268	9.9	8948	4	US-09-221-107-119	Sequence 119, App
25	261.5	9.7	6644	4	US-08-875-435B-5	Sequence 5, Appl
26	254	9.4	6306	1	US-08-195-487-3	Sequence 3, Appl
27	254	9.4	6306	5	PCT-US93-06160-3	Sequence 3, Appl
28	253	9.4	6306	1	US-08-466-390-3	Sequence 3, Appl
29	253	9.4	6306	1	US-08-470-950-3	Sequence 3, Appl
30	253	9.4	6306	1	US-08-467-781-3	Sequence 3, Appl
31	253	9.4	6306	2	US-08-483-924-3	Sequence 3, Appl
32	250.5	9.3	5185	4	US-09-976-594-640	Sequence 640, App
33	248.5	9.2	3824	4	US-09-888-188B-20	Sequence 20, Appl
34	248.5	9.2	3824	4	US-09-291-417D-20	Sequence 20, Appl
35	248.5	9.2	7596	4	US-09-023-655-1463	Sequence 1463, Ap
36	247	9.1	3312	3	US-09-060-410-1	Sequence 1, Appl
37	247	9.1	3312	4	US-09-723-458-1	Sequence 1, Appl
38	246	9.1	5857	4	US-09-220-132-79	Sequence 79, Appl
39	244.5	9.0	6617	4	US-09-976-594-268	Sequence 268, App
40	241.5	8.9	7453	4	US-09-620-312D-248	Sequence 248, App
41	241.5	8.9	7501	4	US-09-620-312D-249	Sequence 249, App
42	240	8.9	8257	4	US-09-595-684B-30	Sequence 30, Appl
43	239	8.8	8503	4	US-09-620-312D-130	Sequence 130, Appl
44	237.5	8.8	4892	4	US-09-620-312D-1065	Sequence 1065, Ap
45	237	8.8	5467	4	US-09-976-594-735	Sequence 735, App
46	235	8.7	3402	4	US-09-252-991A-15560	Sequence 15560, A
47	235	8.7	3606	4	US-09-252-991A-15688	Sequence 15688, A
48	233.5	8.6	6175	4	US-08-875-435B-1	Sequence 1, Appl
49	231	8.5	3702	4	US-09-252-991A-6666	Sequence 6666, Ap
50	230.5	8.5	2934	4	US-09-010-147B-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1
US-08-979-608A-17
; Sequence 17, Application US/08979608A
; Patent No. 6355451

GENERAL INFORMATION:

APPLICANT: Lees, Ann M.

Lees, Robert S.

Law, Simon W.

Arjona, Anibal A.

TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
TREATING ATHEROSCLEROSIS

NUMBER OF SEQUENCES: 42

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/979,608A

FILING DATE: 26-No. 6355451-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/048,547

FILING DATE: 03-JUN-1997

APPLICATION NUMBER: US 60/031,930
 FILING DATE: 27-NOV-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Myers, Louis
 REGISTRATION NUMBER: 35,965
 REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617/542-5070
 TELEFAX: 617/542-8906
 INFORMATION FOR SEQ ID NO: 17:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 4697 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 FEATURE:
 NAME/KEY: Coding Sequence
 LOCATION: 3...1592
 SEQUENCE DESCRIPTION: SEQ ID NO: 17:
 US-08-979-608A-17

Alignment Scores:
 Pred. No.: 9,05e-228 Length: 4697
 Score: 2702.00 Matches: 530
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-10-023-523-8 (1-530) x US-08-979-608A-17 (1-4697)

QY	1	LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln	20
DB	3	AAAGCAGCCAGCAGCAACCGGAAGCAGGACCCGAGCGGAGCCAGGAGCGCCAGCCAG	62
QY	21	AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu	40
DB	63	GGGGCTCTCTGCAATAGAACGCAAGGATCCCGGAGCAGCCAGGCTCTCTGGAAGCGGAG	122
QY	41	GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer	60
DB	123	GGGGCTCAGCCAGCAGACGCTCAGTCTGGGGCCCTTCGTGATGCTCTCTGAGGAGCTGAGC	182
QY	61	ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnGlnGlyGlyProGly	80
DB	183	CGCCAACTGGAAGACATACTGAGCACAATCTGTGGCAATAACACGAGGGGCCCCGGC	242
QY	81	GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLeuSerArgThrTyr	100
DB	243	GAGGATGGGGCAGGGTGAGCGGCTGAACCCAGAGATGAGAACTCCCGGACCTAT	302
QY	101	ValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGluLysGluProSer	120
DB	303	GTGGCAAGGAATGGGAGCCTGAACCAACTCCAGTAGTCTATGAGAGAGAGAACCTCC	362
QY	121	LysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHis	140
DB	363	AAGGGGGATCCAAACACAGAGAGATCCGGCAGAGTGCAGAGGTCGGAGACCGAGACCAT	422
QY	141	ArgArgProGlnGluLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMet	160
DB	423	CGAAGGCCACAGGAGAGAAAACCCAAAGGGTTGGGAGAGGAGATCACGTTCTGATG	482
QY	161	GlnThrLeuAsnThrLeuSerThrProGluLysLeuAlaLeuCysLysLysTyr	180
DB	483	CAGACATTGAATACTCTGAGTACCCAGAGAGAGCTGGCTGCTCTGTGCAAGAGATAT	542
QY	181	AlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLys	200
DB	543	GCTGAATGCTGGAGGAGCACCAGGATTCACAGAGCAGATGAAGCTCTCTACAGAAAAG	602
QY	201	GlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeu	220

DB	603	CAGAGCCAGCTGGTCAAGAGAGAACCACTGCGCGGTGAGCACAGCAAGCGCGCTCCTG	662
QY	221	AlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLys	240
DB	663	GCCCGCAGCAAGCTTGAGAGCTATGCGGTGAGCTGCGAGCGGCACACCGCTCCTCAAG	722
QY	241	GluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHis	260
DB	723	GAAGAAGGTGTGACGCGGCCCGGAGGAGGAGAGCCGACAGAGGTGACCTCGCAC	782
QY	261	PheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSer	280
DB	783	TTCCAGGTGACACTGATGACATTCAGCTGCAGATGGAACACAGACCAATGAGCGCACTCC	842
QY	281	LysLeuArgGlnGlnAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGluGlnTyr	300
DB	843	AAGCTGCGCCAGAGACATGAGCTGGCTGAGAGGCTCAAGAAGCTGATTGACAGTAT	902
QY	301	GluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeu	320
DB	903	GAGCTGCGCGAGGAGCATATCGCAAAAGTCTTCAACACCAAGACCTACAACAGCAGCTG	962
QY	321	ValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGln	340
DB	963	GTGATGCCAAGCTCCAGCAGGCCCGCAGAGATGCTAAAGAGGCGAGAGCGGCACAG	1022
QY	341	ArgGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMet	360
DB	1023	CGGAGAGAGGATTTCTCTGAAAGAGCAGTAGTAGTCCAGAGGATGTGTGAGTGATG	1082
QY	361	LysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGlu	380
DB	1083	AAGCAGCAGAGAGACCCACCTGAAGCAACAGCTTGCCCTATACACAGAGAGTTTGAGGAG	1142
QY	381	PheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGlu	400
DB	1143	TTCCAGAACACACTTTCCAAAAGCAGCAGGATTCACACACATTCAGCAGAGAGATGGAA	1202
QY	401	LysMetThrLysLysLysLysLysLeuGluLysGluThrThrMetTyrArgSerArgTyr	420
DB	1203	AAGATGACTAAGAGAGATCAAGAGCTGGAGAAAGAAACCAACCATGTACCGGTCCCGTGG	1262
QY	421	GluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGlu	440
DB	1263	GAGAGCAGCAACAGCCCTGCTTGAGATGGCTGAGAGAGAAACACCTCCGGGATAAAGAA	1322
QY	441	LeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThr	460
DB	1323	CTGGAGGGCTTGAGGTAATAAATCCAAACGCTGGAGAGCTGTGCCGGGCACTGCAGACA	1382
QY	461	GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlnGlySerLeu	480
DB	1383	GACCGCAATGACCTGAACAGAGGGTACAGGACCTGAGTGTGGTGGCCAGGGCTCCCTC	1442
QY	481	ThrAspSerGlyProGluArgArgProGluGlyProGlyValAlaGlnAlaProSerPro	500
DB	1443	ACTGACAGTGGCCCTGAGAGAGGCCAGAGGGCCCTGGGGCTCAAGACCCAGCTCCCTC	1502
QY	501	ArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThr	520
DB	1503	AGGCTCACAGAGCGCTTGCTTACCAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	1562
QY	521	GlyProGlnGluProThrSerAlaArgAla	530
DB	1563	GGGCTCAAGAGCCACCTCCCGCAGGGCC	1592

RESULT 2

US-09-517-849-17
 ; Sequence 17, Application US/09517849
 ; Patent No. 6605588
 ; GENERAL INFORMATION:
 ; APPLICANT: Lees, Ann M.
 ; ; Lees, Robert S.

Law, Simon W.
 Atjona, Anibal A.
 TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
 BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
 TREATING ATHEROSCLEROSIS
 NUMBER OF SEQUENCES: 42
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/517,849
 FILING DATE: 02-Mar-2000
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/979,608
 FILING DATE: 26-Nov-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Myers, Louis
 REGISTRATION NUMBER: 35,965
 REFERENCE/DOCKET NUMBER: 10797-003001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617/542-5070
 TELEFAX: 617/542-8906
 INFORMATION FOR SEQ ID NO: 17:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 4697 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 FEATURE:
 NAME/KEY: Coding Sequence
 LOCATION: 3...1592
 SEQUENCE DESCRIPTION: SEQ ID NO: 17:
 US-09-517-849-17

Alignment Scores:
 Pred. No.: 9.05e-228 Length: 4697
 Score: 2702.00 Matches: 530
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-10-023-523-8 (1-530) x US-09-517-849-17 (1-4697)
 Qy 1 LysSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 20
 Db 3 AANAAGCAGCCAGGACCAACCGAAGCAGGACCCGAGGAGCCAGGAGCGCCAGCAG 62
 Qy 21 AlalaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40
 Db 63 GCGGCTCCTGAGTAGAAGCAGAGAGGTCGCGGAGCAGCAGGCTCTCTCGAAGCCGAG 122
 Qy 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 60
 Db 123 GGGGCTCAGCAGCAGCAGGCTCAGTCTGGGGCCCTTCGTGATGCTCTGAGGAGCTGAGC 182
 Qy 61 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnGlnGlyGlyProGly 80
 Db 183 CGCCAACCTGGAAGACATCTGAGCAGCATCTGTGTGACAAATACACAGGGGGGCCCCGGC 242
 Qy 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
 Db 243 GAGGATGGGGCAGAGGAGTGTAGCCGCTGAACCCGAAAGATGCAGAGAAAGTCCCGACCTAT 302

Qy 101 ValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGluLysGluProSer 120
 Db 303 GTGGCAAGGAATGGGAGCCTGAACCAACTCCAGTAGTCTATGAGAGAGAGAACCCCTCC 362
 Qy 121 LysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHis 140
 Db 363 AAGGGGATCCAAACACAGAGAGATCCGCGAGGTGACGAGGTGCGAGACCGAGACCAT 422
 Qy 141 ArgArgProGlnGluLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMet 160
 Db 423 CGAAGGCCACAGGAGAGAAAAAGCCAAAGGTTTGGGGAGGAGATCACGTTGCTGATG 482
 Qy 161 GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLysTyr 180
 Db 483 CAGACATTGATCTCTGATACCCAGAGAGAGTGGCTGCTCTGTGCAAGAATAT 542
 Qy 181 AlaGluLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLys 200
 Db 543 GCTGAACCTGCTGGAGGAGCACCGGAATTCAGAAAGCAGATGAAGCTCTCACAGAAAAG 602
 Qy 201 GlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeu 220
 Db 603 CAGAGCCAGCTGGTGCAAGAGAGACCCCTGCGCGGTGAGCACAGCAAGGCCGCTCTG 662
 Qy 221 AlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLysLys 240
 Db 663 GCCCGCAGCAAGCTTGAGAGCCTATGCGTGTGAGCGGCGCAACCCGCTCCCTCAAG 722
 Qy 241 GluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSerHis 260
 Db 723 GAAGAAGGTGTGCAAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 782
 Qy 261 PheGlnValThrLeuAsnAspIleGlnLeuMetGluGlnHisAsnGluArgAsnSer 280
 Db 783 TTCAGGTGACACTGAATGATCATTCAGCTGCAGTGGAGACAGCACATGAGCGCACTCC 842
 Qy 281 LysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuIleGlnTyr 300
 Db 843 AAGCTGCGCCAAAGAACATGAGGTGCTGAGAGGCTCAAGAGGCTCAATTGAGCAGTAT 902
 Qy 301 GluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeu 320
 Db 903 GAGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 962
 Qy 321 ValAspAlaLysLeuGlnGlnAlaGlnMetLeuLysGluAlaGluLysGlnHisGln 340
 Db 963 GTGGATGCCAAGCTCCAGCAGGCCCCAGGAGATGCTTAAAGGAGGAGGAGGAGGAGGAG 1022
 Qy 341 ArgGluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMet 360
 Db 1023 CGGGAGAAAGGATTTCTCTCTGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1082
 Qy 361 LysGlnGlnGluThrHisLysLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGlu 380
 Db 1083 AAGCAGCAGAGAGACCCACCTGAGCAACAGCTTGCCTTATACACAGAGAGTTTCAGCAG 1142
 Qy 381 PheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGlnMetGlu 400
 Db 1143 TTCAGAACACACTTTCACAAAGCAGCAGGAGTATTCACCACATTCACAGCAGGAGTGA 1202
 Qy 401 LysMetThrLysLysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArgTyr 420
 Db 1203 AAGATGACTTAAAGAAATCAAGAGCTGGAGAAAGAAACCAACCATGTACCGTCCCGTGG 1262
 Qy 421 GluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGlu 440
 Db 1263 GAGAGCAGCAACAGGCCCTGCTTGCAGATGGCTGAGGAGAGAGAGAGAGAGAGAGAG 1322
 Qy 441 LeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThr 460
 Db 1323 CTGGAGGGGCTCAGGATAAAATCCAAACGCTGGAGAGAGCTGTGCGGGCAGCTGAGACA 1382
 Qy 461 GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlySerLeu 480

Db 1383 GAGCGCATGACCTGAACAAGAGGTACAGGACCTGAGTGTGTCGCGAGGCTCCCTC 1442
Qy 481 ThrAspSerGlyProGluArgProGluGlyProGluGlyProGlyAlaGlnAlaProSerPro 500
Db 1443 ACTGACAGTGGCCCTGAGAGGAGGCCAGAGGGGCTCGGGCTCAAGCACCCAGCTCCCCC 1502
Qy 501 ArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThr 520
Db 1503 AGGTCACAGAGGCGCTTGCTACCCAGAGCACCAGGACGACAGAGCATCAGGCCAGACT 1562
Qy 521 GlyProGlnGluProThrSerAlaArgAla 530
Db 1563 GGGGCTCAAGAGCCACCTCGCCAGGGCC 1592

RESULT 3

US-09-616-289-17
; Sequence 17, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 4697
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(1592)
US-09-616-289-17

Alignment Scores:

Pred. No.: 9,05e-228 Length: 4697
Score: 2702.00 Matches: 530
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-023-523-8 (1-530) x US-09-616-289-17 (1-4697)

Qy 1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 20
Db 3 AAAAGCAGCCAGGACCAACCGGAAGCAGGACCGAGGAGGCCAGGAGGCCAGGCCAG 62
Qy 21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerSerAlaProArgLysProGlu 40
Db 63 GCGGCTCCCTGAGTAGAAGCAGAGGCTCCCGGAGCAGGCTCCCTCGGAAGCCGGAG 122
Qy 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluLeuSer 60
Db 123 GGGGCTCAAGCCAGGACGCTCAGTCTGGGGCCCTTCGTGATGTCTCTGAGGAGCTGAGC 182
Qy 61 ArgGlnLeuGluAspIleuSerThrTyrCysValAspAsnGlnGlyClyProGly 80

Db 183 CGCCAACTGGAAAGACATACTGAGCACATACCTGTGTGGACAATAACACGGGGGGCCCCGGC 242
Qy 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
Db 243 GAGGATGGGCGACAGGGTGAGCCGGCTGAACCCGAAGATGACAGAGAGTCCCGGACCTAT 302
Qy 101 ValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGluLysGluProSer 120
Db 303 GTGGCAAGAAATGGGAGCGCTGAACCAACTCCAGTAGTCTATGGAGAGAAAGAACCCCTCC 362
Qy 121 LysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHis 140
Db 363 AAGGGGGATCCAAACACAGAGAGATCCGCGAGAGTGCAGAGGTCGGAGACCGAGACCAT 422
Qy 141 ArgArgProGlnGlnGlyLysLysAlaLysGlyLysGlyLysGluIleThrLeuLeuMet 160
Db 423 CGAAGGCCACAGGAGAGAAAGAAAGCCAAAGGGTTTGGGGAAGGAGATCACCTTCTGATG 482
Qy 161 GlnThrLeuAsnThrLeuSerThrProGluGlyLysLeuAlaAlaLeuCysLysLysTyr 180
Db 483 CAGACATTGAATCTCTGAGTACCCAGAGAGAGAGCTGGCTGCTCTGTGCAAGAGTAT 542
Qy 181 AlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMethLysLeuLeuGlnLysLys 200
Db 543 GCTGAATCTCTGGAGGAGCACCGGAATTCACAGAAGCAGATGAAGCTCCTACAGAAAAG 602
Qy 201 GlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeu 220
Db 603 CAGAGCCAGCTGTGTCAAGAGAGACCACTCGCGGTGAGACAGCAGCAAGCGCTCTG 662
Qy 221 AlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLysLys 240
Db 663 GCGCGCAGCAAGCTTGAGAGCTATGCCGTGAGCTGCAGCGGCACAAACCGCTCCTCAAG 722
Qy 241 GluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGlyLysValThrSerHis 260
Db 723 GAAGAAGGTGTGACGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 782
Qy 261 PheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSer 280
Db 783 TTCCAGGTGACACATGATGACATTCAGTGCAGATGGAAACAGCACATGAGCCCACTCC 842
Qy 281 LysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLysLeuIleGluGlnTyr 300
Db 843 AAGCTGCGCCCAAGAGAACATGGAGCTGGCTGAGAGGCTCAAGAAGCTGATTGAGCAGTAT 902
Qy 301 GluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeu 320
Db 903 GAGCTGCGGAGGAGCATATCCCAAAAGTCTTCAACACAGAGACCTTCAACAGCAGCTG 962
Qy 321 ValAspAlaLysLeuGlnGlnAlaGlnGluMetLysLysGluAlaGluArgHisGln 340
Db 963 GTGGATGCCAAGTCCAGCAGGCGCCAGGAGATGCTAAAGAGGAGCAGAGGAGCGCCAGCAG 1022
Qy 341 ArgGluLysAspPheLeuLeuLysGluAlaValCysSerGlnArgMetCysGluLeuMet 360
Db 1023 CGGAGAGAGGATTTCTCTGAAAGAGGAGTACAGTCCAGAGGAGTGTGTGAGCTGATG 1082
Qy 361 LysGlnGlnGluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGlu 380
Db 1083 AAGCAGCAAGAGAGACCCACCTGAGCAACACAGCTTGCCCTATACACAGAGAGTTTGGAGG 1142
Qy 381 PheGlnSerThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGlu 400
Db 1143 TTCCAGAACACACTTTCCAAAAGCAGCAGGATATCCACCATTCAGCAGGAGATGGAA 1202
Qy 401 LysMetThrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 420
Db 1203 AAGATGACTAAGAGATCAAGAGAGTGGAGAAAGAACACCATGTTACCGTCCCGGTGG 1262
Qy 421 GluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGlu 440
Db 1263 GAGAGCAGCAACAGGCGCTTGTGATGGCTGAGGAGAGAAACAGTCCGGGATAAGAA 1322

QY 441 LeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThr 460
Db 1323 CTGAGGGCCCTGCGAGTAAATAATCCACGGCTGAGAGAGCTGTGCGGGGCACTGCAGACA 1382
QY 461 GlnArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeu 480
Db 1383 GAGCGCAATGACCTGAACAGAGGGTACAGGACTGTGCTGGTGGCAGGGCTCCCTC 1442
QY 481 ThrAspSerGlyProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerPro 500
Db 1443 ACTGACAGTGGCCCTGAGAGAGAGCCAGAGGGGCTTGGGGCTCAAGCAGCCAGCTCCCC 1502
QY 501 ArgValThrGluAlaProCysTyrProGlyValaProSerThrGluAlaSerGlyGlnThr 520
Db 1503 AGGTCACAGAGGGCTTGTCTACCCAGAGGCCAGCCAGGACACAGAGCATCAGGCCAGCT 1562
QY 521 GlyProGlnGluProThrSerAlaArgAla 530
Db 1563 GGGGCTCAAGAGCCAGCTCCGCGAGGGCC 1592

RESULT 4

US-09-616-289-46
; Sequence 46, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 1638
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1638)
US-09-616-289-46

Alignment Scores:
Pred. No.: 1,19e-227 Length: 1638
Score: 2693.00 Matches: 529
Percent Similarity: 99.81% Conservative: 0
Best Local Similarity: 99.81% Mismatches: 1
Query Match: 99.67% Indels: 0
Gaps: 4
DB:

US-10-023-523-8 (1-530) x US-09-616-289-46 (1-1638)

QY 1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnArgProSerGln 20
Db 49 AAAAGCAGCCAGGACCAACCGAAGAGGAGCCAGGAGGCCAGGAGGCCAGGCCAG 108
QY 21 AlaAlaProAlaValGluAlaGlyProGlySerGlnAlaProArgLysProGlu 40
Db 109 GCGGCTCTCGAGTAGAGCAGAAAGGTCCCGGACAGCAGGCTCTCTGAGAGCCGAG 168

QY 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 60
Db 169 GGTGCTCAAGCCAGAACGGCTCAGTCTGGGGGCCCTTCGTGATGTCTCTCAGAGAGTGA 228
QY 61 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnGlnGlyGlyProGly 80
Db 229 CGCCACTGGAGACATACCTAGCACATACCTGTGTGGACAATAACAGGGGGGCCCGGC 288
QY 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
Db 289 GAGGATCGGGCAGAGGGTGAAGCGGCTGAACCGGAAGATGCCAGAGAGTCCCGGACCTAT 348
QY 101 ValAlaArgAsnGlyGluProGluProThrProValValTyrGlyLysGluProSer 120
Db 349 GTGGCAAGGATGGGAGGCTGAACCACTCCAGTAGTCAATGGAGAGAGAACCTCC 408
QY 121 LysGlyAspProAsnThrGluLysIleArgGlnSerAspGluValGlyAspArgAspHis 140
Db 409 AAGGGGATCCAAACACAGAGAGATCCGGCAGAGTGCAGAGTCCGAGACCGAGACCAT 468
QY 141 ArgArgProGlnGluLysLysAlaLysGlyLeuGlyLysGluLysGluLysLeuMet 160
Db 469 CGRAGGCCACAGGAGAGAAAGCCAGGGTTTGGTAAGGAGATCACGTTGCTGATG 528
QY 161 GlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysTyr 180
Db 529 CAGACATTGAATATCTGTGATACCCAGAGAGAGCTGGCTCTGTGCAAGAGAT 588
QY 181 AlaGluLeuLeuGluGlnHisArgAsnSerGlnLysGlnMetLysLeuGlnLysLys 200
Db 589 GCTGAATGCTGGAGAGGACCGGATTCACAGACAGATGAAGCTCTACAGAAAAG 648
QY 201 GlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeu 220
Db 649 CAGAGCAGCTGTGCAAGAGAGAGACCACTGCGCGGTGAGCAGCAGAGGCGGCTCTG 708
QY 221 AlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLys 240
Db 709 GCCCGCAGCAGCTTGAGAGCTATGCCGTGAGCTGCAGCGGCACACCGCTCCCTCAG 768
QY 241 GluGlyValGlnArgAlaArgGluGluGluGluGluGluGluGluGluGluGluGlu 260
Db 769 GAAGAAGGTGTGAGCGGGCCCGGAGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 828
QY 261 PheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSer 280
Db 829 TTCAGGTGACACTGAATGACATTCAGTGCAGTGAACAGACACATCAGCGCACTCC 888
QY 281 LysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLysLysLysLys 300
Db 889 AAGCTGGCCCAAGAGAACATGAGCTGGCTGAGAGGCTCAAGAGAGCTGATTGAGCAGTAT 948
QY 301 GluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeu 320
Db 949 GAGCTGGCGAGGAGCATATCGACAAAGTCTTCAACACAGAGGACCTACACAGCAGCTG 1008
QY 321 ValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGln 340
Db 1009 GTGGATCCCAAGTCCAGCAGCGCCAGGAGATGCTAAAGAGGAGCAGAGAGCGGACCCAG 1068
QY 341 ArgGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMet 360
Db 1069 CGGGAGAGAGATTTCTCTGAAAAGGCGAGTAGACTCCAGAGGATGTGTGAGCTGATG 1128
QY 361 LysGlnGlnGlnThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGlu 380
Db 1129 AAGCAGCAAGAGAGACCCACCTGAAGAGCAACAGCTTGCCTATACACAGAGAGATTGAGGAG 1188
QY 381 PheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPhenylsGlnGluMetGlu 400
Db 1189 TTCAGAGACACACTTCCAAAGACGAGGATATTCCACCATTCAGAGAGAGATGGA 1248
QY 401 LysMetThrThrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 420

Db 1249 AAGATGACTAAGAGATCAAGAGCTGGAGAAAGAAACCACTGATACCGTCCCGTGG 1308
QY 421 GluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGlu 440
Db 1309 GAGAGCAGCAACAGGCGCTCTTGATGCTGAGGAGAAACAGTCCGGGATAAGAA 1368
QY 441 LeuGluGlyLeuGlnValLysLeuGlnArgLeuGluLysLeuCysArgAlaLeuGlnThr 460
Db 1369 CTGGAGGCGCTCGCAGTGAATAAATCAACGCTGGAGAGCTGTGCCGGGCACTGCACACA 1428
QY 461 GluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlnGlySerLeu 480
Db 1429 GAGCGCAATGACCTGAACAGAGGGTACAGACCTGTGCTGCTGCGCAGGCTCCCTC 1488
QY 481 ThrAspSerGlyProGluArgArgProGluGlyProGluGlyProGluGlnAlaProSerPro 500
Db 1489 ACTGACAGTGGCTCGAGAGAGGAGCCAGAGGGGCTGGGGCTCAAGCACCCAGCTCCCC 1548
QY 501 ArgValThrGluAlaProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThr 520
Db 1549 AGGGTCAAGAGAGCGCTTGCTCCAGGAGCAGCCAGAGCAGAGCATCAGGCCAGACT 1608
QY 521 GlyProGlnGluProThrSerAlaArgAla 530
Db 1609 GGGCTCAAGAGCCACCTCGCCAGCGGCC 1638

RESULT 5

US-08-979-608A-14

; Sequence 14, Application US/08979608A

; Patent No. 6355451

; GENERAL INFORMATION:

; APPLICANT: Lees, Ann M.

; Lees, Robert S.

; Law, Simon W.

; Arizona, Anibal A.

; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN

; BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND

; TREATING ATHEROSCLEROSIS

; NUMBER OF SEQUENCES: 42

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/979,608A

; FILING DATE: 26-Jun-1997

; PRIORITY DATE: 26-Jun-1997

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: US 60/048,547

; FILING DATE: 03-JUN-1997

; APPLICATION NUMBER: US 60/031,930

; FILING DATE: 27-NOV-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Myers, Louis

; REGISTRATION NUMBER: 35,965

; REFERENCE/DOCKET NUMBER: 10797-002001 (formerly 3983/59818)

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617/542-5070

; TELEFAX: 617/542-8906

; INFORMATION FOR SEQ ID NO: 14:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4722 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 61...1731
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-08-979-608A-14
Alignment Scores:
Pred. No.: 5,91e-207 Length: 4722
Score: 2465.50 Matches: 492
Percent Similarity: 92.99% Conservative: 12
Best Local Similarity: 90.77% Mismatches: 25
Query Match: 91.25% Indels: 13
DB: 4 Gaps: 3
US-10-023-523-8 (1-530) x US-08-979-608A-14 (1-4722)
QY 1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyValaGlnGluArgProSerGln 20
Db 109 AAAAGCAGCCCGGACAGCGGAGCAGGAGCGGAGGAGCCAGGGCGGCGCCGCGG 168
QY 21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40
Db 169 CCGCCCCCCCCCGAGAGCCGGAAGGT---GCCAGCAGCCAGGCTCCCGGAGCGCGAG 225
QY 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSer 60
Db 226 GGGGCTCAAGCCAAACTGCTCAGCCCTGGGGCGCTCTGTGATGTCTCTGAGGAGTGA 285
QY 61 ArgGlnLeuGluAspLysLeuSerThrTyrCysValAspAsnAsnGlnGlyProGly 80
Db 286 CGCCAGTTGGAAGACATACACTAGTACATCTGTGTGGCAACAACAGCGGGCGCCGCGGT 345
QY 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
Db 346 GAGGATGGGTCCAGGGTGAGCCCCCTGAACCTGAAGATGCAGAGAGTCTCGCCCTAT 405
QY 101 ValAlaArgAsnGlyGluProGluPro---ThrProValValTyrGlyGluLysGluPro 119
Db 406 GTGCAAGGAATGGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 465
QY 120 SerLysGlyAspProAsnThrGluGluLysLeuArgGlnSerAspGluValGlyAspArgAsp 139
Db 466 TCCAGGCGAGCGCGGAGCGGAGAGATCCCGAGCGGAGCGGAGCGGAGCGGAGCGGAG 525
QY 140 HisArgArgProGlnGluLysLysAlaLysGlyLeuGlyLysGluLysLeuLysLeu 159
Db 526 CACCGAGGCGCACAGGAAAGAGAGGCGGAGGCTCTGGGAAAGAGAGATCACTCGCTGTG 585
QY 160 MetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuLysLysLys 179
Db 586 ATGAGACACTGACACGCTGAGCAGCCCGAGGAGAGAGTGGCGGCTCTGTGCAAGAG 645
QY 180 TyrAlaGluLeuLeuGluHisArgAsnSerGlnLysGlnMetLysLeuLysLys 199
Db 646 TATCGCGAATCTGCTCGAGGAGCAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 705
QY 200 LysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaVal 219
Db 706 AGCAGAGCGGAGCTGTGAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 765
QY 220 LeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeu 239
Db 766 CTGCGCCGAGCAAGCTCGAGAGCGCTGTGCCGGGAGCTGACAGCGGCGCAACCGCTCG 825
QY 240 LysGluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSer 259
Db 826 AAGGAGAGAGGTGTGAGGAGCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 885
QY 260 HisPheGlnValThrLeuAsnAspLysGlnLeuGlnMetGluGlnHisAsnGluArgAsn 279
Db 886 CACTTCCAGATGACGCTCAACGACATTCAGTGTGAGAGTGGAGAGCAGCAGCAGCGCAC 945
QY 280 SerLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLeuLysLeuLys 299

Db 946 TCCAGAGCTGCGCAGAGAAATGAGAGTGGCGGCGCTCAAGAGCTGATTGAGCAG 1005
Qy 300 TyrGluLeuArgGluGluHisLeuAspLysValPheLysHisLysAspLeuGluGlnGln 319
Db 1006 TACGAGCTGCGAGAGAGACATCGAAGAGTCTCAACACACAGAGATCTCCAGCAGCAG 1065
Qy 320 LeuValAspAlaLysLeuGlnGlnAlaGlnMetLeuLysGluAlaGluGluArgHis 339
Db 1066 CTGTGGAGCGGCAAGCTCCAGCAGCGCCAGGAGATGCTGAAGGAGGCGAGAGCGGCGAC 1125
Qy 340 GlnArgGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeu 359
Db 1126 CAGCGGAGAGAGACTTCTCTGAGGAGCGCTGGAGTCCAGAGATGTCGAGCTG 1185
Qy 360 MetLysGlnGlnThrHisLeuLysGlnGlnLeuAlaLeuLysThrGluLysPheGlu 379
Db 1186 ATGAAGCAACAGGAGACCCACCTGAAGCAGCAGCTTGCCTATACACAGAGAGTTTGA 1245
Qy 380 GluPheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMet 399
Db 1246 GAGTTCAGAACACTTCTTCCAAAGCAGCAGGAGTTCACCACTTCAACAGGAATG 1305
Qy 400 GluLysMetThrLysLysLysLeuGluLysGluThrThrMetTyrArgSerArg 419
Db 1306 GAAAGATGACAAAGAGATCAAGAGCTGAGAAAGAGACCACTGTCCTCCCGG 1365
Qy 420 TrpGluSerSerAsnLysAlaLeuLeuGluMetAlaGluLysThrValArgAspLys 439
Db 1366 TGGAGAGCAGCAACAGCGCTGCTGAGATGCTGAGAGAGAAACACTCCGGGACAA 1425
Qy 440 GluLeuGluGlyLeuGlnValLysLysGlnArgLeuGluLysLeuCysArgAlaLeuGln 459
Db 1426 GAGTGGAGGCTTGCAGTGAAATCAGCGGCTGGAGAGCTGTCGGGCACTGCAG 1485
Qy 460 ThrGluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlnGlySer 479
Db 1486 ACAGAGCGCAATGACCTGAACAGAGGCTGAGGACCTGAGTGGCGGTGGCCAGGCGCCC 1545
Qy 480 LeuThrAspSerGlyProGluArgPro----- 489
Db 1546 GTCTCCGAGCAGCGTCTGAGCGGAGCGCAGCGCCGCCACCTCCAGAGGAGCAGGGT 1605
Qy 490 ---GluGlyProGlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyr 508
Db 1606 GTCAGAGGCGCGCGGCTCAAGTACCCCACTCTCCAGGCGCCACAGAGCTTCTGCTGC 1665
Qy 509 ProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAla 528
Db 1666 GCAGGTGCACCCAGCAGAGGATCATCGCCAGACAGCGGCCCGCCAGGAGCCCACTGCC 1725
Qy 529 ArgAla 530
Db 1726 ACTGCC 1731

RESULT 6

US-09-517-849-14
: Sequence 14, Application US/09517849
: Patent No. 6605588
: GENERAL INFORMATION:
: APPLICANT: Lees, Ann M.
: Lees, Robert S.
: Law, Simon W.
: Arjona, Anibal A.
: TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
: BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
: TREATING ATHEROSCLEROSIS
: NUMBER OF SEQUENCES: 42
: CORRESPONDENCE ADDRESS:
: ADDRESSSEE: Fish & Richardson P. C.
: STREET: 225 Franklin Street
: CITY: Boston
: STATE: MA

COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/517,849
FILING DATE: 02-Mar-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/979,608
FILING DATE: 26-NOV-1997
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Louis
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: 10797-003001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 4722 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 61...1731
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-517-849-14

Alignment Scores:
Pred. No.: 5,91e-207 Length: 4722
Score: 2465.50 Matches: 492
Percent Similarity: 92.99% Conservative: 12
Best Local Similarity: 90.77% Mismatches: 25
Query Match: 91.25% Indels: 13
DB: 4 Gaps: 3

US-10-023-523-8 (1-530) x US-09-517-849-14 (1-4722)

Qy 1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyValaGlnGluArgProSerGln 20
Db 109 AAAAGCAGCCCGGACAGCCGAGCAGAGCGGAGGAGCCAGGGGCGGCGCGCGCG 168
Qy 21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40
Db 169 CCGGCGCCCGCGCGAGAGCCGAAAGT---GCCAGCAGCCAGGCTCCCGGAGGCGGAG 225
Qy 41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluLeuSer 60
Db 226 GGGGCTCAAGCCAAACTCTCAGCCTGGGGCGCTCTGTGATGTCTCTGAGGAGCTGAGC 285
Qy 61 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyProGly 80
Db 286 CGCCAGTTGGAAGACATCTAGTACATCTGTGGACACACACAGCGGCGCGCGCGGT 345
Qy 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
Db 346 GAGGATGGGGTCCAGGGTTCAGCCCTGAACCTGAAGATGCAGAGAGTCTCGCGCCTAT 405
Qy 101 ValAlaArgAsnGlyGluProGluPro---ThrProValValTyrGlyLysGluPro 119
Db 406 GTGGCAAGAAATGGGAGCGCGGAGCGGCGGAGCCAGCCAGTGTCAATGGCGAGAGAGACC 465
Qy 120 SerLysGlyAspProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAsp 139
Db 466 TCCAAGGCAGAGCCCGGACGAGAGATCCCGAGCAGCGATGAGTCCGAGACCGAGAC 525
Qy 140 HisArgArgProGlnGluLysLysLysAlaLysGlyLeuGlyLysGluLysLeuLeu 159
Db 526 CACCGAGGCGCCACAGAGAAAGAGGCGGCGGAGGCGGCGGAGAGATACCGTCTGCTG 585

160 MetGlnThrLeuAenThrLeuSerThrProGluGlnLysLeuAlaLeuCysLysLys 179
Db ATGAGACACTGAACACGCTGAGCACCACCCAGAGAGAGCTGGCGGCTGTGTGCAAGAAG 645
180 TyrAlaGluLeuLeuGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLys 199
Db TATCGGGAACCTGCTCGAGGAGCACCAGAACTCCGAGAGAGCAGATGAGCTGTGCGAAG 705
200 LysGlnSerGlnLeuValGlnGlnLysAspHisLeuArgGlnGluHisSerLysAlaVal 219
Db AAGCAGAGCCAGCTGGTGCAGCAGACCCAGCACCCTGCTGCGAGCAGCAGCAAGGCCATC 765
220 LeuAlaArgSerLysLeuLeuSerLysLeuCysArgGlnLeuGlnArgHisAsnArgSerLeu 239
Db CTGCGCCGACAGCTGAGAGGCTGTCGCGGAGCTGCGAGGAGCAGCAACCCGCTGCTGCTC 825
240 LysGluGluGluValGlnArgAlaArgGlnGluLysArgLysGlnValThrSer 259
Db AAGNAGAGAGTGTGACAGGAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 885
260 HisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGlnHisAsnGluArgAsn 279
Db CACTTCCAGATGAGCTCAACGACATTCAGCTGCAGATGGAGGAGCAGCAGCAGCGCAAC 945
280 SerLysLeuArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 299
Db TCCAGAGCTGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1005
300 TyrGluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGln 319
Db TACGAGCTGCGAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1065
320 LeuValAspAlaLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 339
Db CTGGTGGAGCGGCAAGCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1125
340 GlnArgGluLysAspPheLeuLysGlnAlaValGluSerGlnArgMetCysGluLeu 359
Db CAGCGGAGAGAGGAGCTTCTCTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1185
360 MetLysGlnGlnGlnThrHisLysLysGlnGlnLeuAlaLeuThrGluLysPheGlu 379
Db ATGAAGCAACAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1245
380 GluPheGlnAsnThrLeuSerLysSerGluValPheThrThrPheLysGlnGluMet 399
Db GAGTTCCAGAACACTCTTCCAAAGAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1305
400 GluLysMetThrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 419
Db GAAAAGATGACAAAGAGATCAAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1365
420 TrpGluSerSerAsnLysAlaLeuLeuGlnMetAlaGluLysThrValArgAspLys 439
Db TCGGAGAGCAGCAACAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1425
440 GluLeuGluGluLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 459
Db GAGCTGGAAGGCTGCGAGTGAAATCCAGCGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1485
460 ThrGluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlnGlySer 479
Db ACAGAGCGCAATGACTGTACAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1545
480 LeuThrAspSerGlyProGluArgArgPro----- 489
1546 GTCCTCCGAGAGGCTCTGAGCGGAGGAGCAGAGCGCGCCGCCACCTCCAGAGGAGCAGGAGT 1605
490 ----GluGlyProGlyAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyr 508
1606 GTCGAGGGGCGCGGGGCTCAAGTACCAACTCTCCAGAGGAGCAGAGAGCTTCTGCTGTCG 1665

509 ProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAla 528
Db GCAGGTGACCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1725
529 ArgAla 530
Db ACTGCC 1731

RESULT 7
US-09-616-289-14
; Sequence 14, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arizona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 4722
; TYPE: DNA
; ORGANISM: Oryctolagus cuniculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)....(1731)
US-09-616-289-14

Alignment Scores:
Pred. No.: 5,916-207 Length: 4722
Score: 2465.50 Matches: 492
Percent Similarity: 92.99% Conservative: 12
Best Local Similarity: 90.77% Mismatches: 25
Query Match: 91.25% Indels: 13
DB: 3 Gaps: 3

US-10-023-523-8 (1-530) x US-09-616-289-14 (1-4722)

1 LysSerSerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGln 20
Db AAAAGCAGCCGCGAGCAGCCGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 168
21 AlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlu 40
Db CCGGCGCCCGCGAGAGAGCCGAGGT---GCCAGCAGCCAGGCTCCCGGAGGAGGAGGAGGAG 225
41 GlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluLeuLeuSer 50
Db GGGGCTCAAGCCAAACTCTCAGCCTGGGGGCTCTGTGATGTCTCTGAGGAGCTGAGC 285
61 ArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnGlnGlyGlyProGly 80
Db CGCAGTTGAGAGACATCTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 345
81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGlnLysSerArgThrTyr 100
Db GAGGATGGGTTCAGGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 405
101 ValAlaArgAsnGlyGluProGluPro---ThrProValValTyrGlyGluLysGluPro 119

Db 406 GTGCAAGGAATGGGAGCGGAGCGGCGGACCCCACTAGTCAATGGCGAGAGGAGACC 465
Qy 120 SerLysGlyAspProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArgAsp 139
Db 466 TCAAGGACAGAGCGGCGGACCGGAGAGATCCGACGAGCGATGAGTCCGAGACCGAGAC 525
Qy 140 HisArgArgProGlnGluLysLysAlaLysGlyLeuGlyLysGluLeuThrLeuLeu 159
Db 526 CACCGAGGCCACAGGAAGAGAGAGGCCCAAGGGCTCGGAAGGAGATCAGCGTGCTG 585
Qy 160 MetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCysLysLys 179
Db 586 ATGCAGACACTGAACGCTGAGCACCCGAGAGAGAGCTGGCGGCTCTGTCAAGAAAG 645
Qy 180 TyrAlaGluLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLys 199
Db 646 TATGCGGAATGCTCGAGGAGACCGGAACTCCGAGAGCAGATGAAGCTGCTGCAGAG 705
Qy 200 LysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAlaVal 219
Db 706 AAGCAGAGCCAGCTGTGTCAGGAGAGAGACCACTCGCTGGCGAGCACAGCAAGGCCATC 765
Qy 220 LeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeu 239
Db 766 CTGGCCGCGAGCAAGCTCGAGAGCGCTGTGCGGAGCTCGAGCGGCACCAACCGCTCGTC 825
Qy 240 LysGluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSer 259
Db 826 AAGGAAGAAGTGTGACGAGCGCCGAGAGAGGAGGAGAGCGCAAGGAGGTGACGTCA 885
Qy 260 HisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsn 279
Db 886 CACTTCCAGATGAGCTCAACGACATTCAGCTCAGATGAGCAGCAGCAACGAGCGCAAC 945
Qy 280 SerLysLeuArgGlnGlnAsnMetGluLeuAlaGluArgLeuLysLysLeuLeuGln 299
Db 946 TCCAAGCTGCCAGAGAGACATGGAGCTGCCGAGCGGCTCAAGAGCTGATTGAGCAG 1005
Qy 300 TyrGluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGln 319
Db 1006 TAGCAGCTGCGAGAGAGACATCGAAGAGTCTTCAACACAGAGATCTGACAGCAGCAG 1065
Qy 320 LeuValAspAlaLysLeuGlnGlnAlaGlnGluMetLysGluAlaGluArgHis 339
Db 1066 CTGGTGACCCCAAGCTTCCAGAGGCGCCAGGATGCTGAAGGAGGAGGAGCGGCGAC 1125
Qy 340 GlnArgGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeu 359
Db 1126 CAGCGGAGAGAGACTTCTCTGAGAGGCGCTGGAGTCCGAGAGGATGCGAGCTG 1185
Qy 360 MetLysGlnGlnThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGlu 379
Db 1186 ATGAAGCAACAGAGAGACCCACCTGGAAGCAGCAGCTTGCCCTATACACAGAGATTTGAG 1245
Qy 380 GluPheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMet 399
Db 1246 GAGTTCAGAACACTCTTTCAGAGAGAGAGGCTTACACATTCAGACAGGAAATG 1305
Qy 400 GluLysMetThrLysLysLysLysLysLeuGluLysGluThrThrMetTyrArgSerArg 419
Db 1306 GAAAGATGACAAAGAGATCAAGAGCTGGAGAAAGAGAGACCATGATACCGTTCCCGG 1365
Qy 420 TrpGluSerSerAsnLysAlaLeuLeuGluMetAlaGluLysThrValArgAspLys 439
Db 1366 TGGAGAGACGACACAGGCGCTGCTGAGATGGCTGAGAGAAACACTCCGGGCAAA 1425
Qy 440 GluLeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGln 459
Db 1426 GAGCTGGAAGGCTGCGAGGTGAAATCCAGCGCTGGAGAGCTGTGCGGCGACTGCAG 1485
Qy 460 ThrGluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGlyGlnGlySer 479

Db 1486 ACAGAGCGCAATGACCTGAACAAGAGGGTGCAGGACCTGAGTCCGCTGGCCAGGGCCCC 1545
Qy 480 LeuThrAspSerGlyProGluArgArgPro 489
Db 1546 GTCTCCGACAGCGGCTCTGAGCGGAGCCAGAGCCGCCACACCTCCAAAGAGAGAGGT 1605
Qy 490 ---GluGlyProGlyValaGlnAlaProSerProArgValThrGluAlaProCysTyr 508
Db 1606 GTCAGAGGGCCCGGGGCTCAAGTACCACTCTCAAGGGCCACAGAGCGCTTCTGTCTGC 1665
Qy 509 ProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGluProThrSerAla 528
Db 1666 GCAGGTGCACCCAGCAGACAGAGGCATCAGGCCAGAGGGCCCCAGAGCCCACTGCC 1725
Qy 529 ArgAla 530
Db 1726 ACTGCC 1731

RESULT 8
PCT-US94-01101-1/c
; Sequence 1, Application PC/TUS9401101
; GENERAL INFORMATION:
; APPLICANT:
; APPLICANT:
; APPLICANT:
; APPLICANT:
; APPLICANT:
; TITLE OF INVENTION: HIGH MOLECULAR WEIGHT B-CELL GROWTH
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSES: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/01101
; FILING DATE: FILED HEREWITH
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/005,156
; FILING DATE: 15-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: GATES, EDWARD R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: B0819/7000WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/720-3500
; TELEFAX: 617/720-2441
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1854 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: HOMO SAPIENS
; CELL TYPE: LYMPHOMA
PCT-US94-01101-1

Alignment Scores:
Pred. No.: 3,98e-131 Length: 1854
Score: 1597.00 Matches: 334

Percent Similarity: 97.66% Conservative: 0
Best Local Similarity: 97.66% Mismatches: 5
Query Match: 59.10% Indels: 5
DB: Gaps: 0

US-10-023-523-8 (1-530) x PCT-US94-01101-1 (1-1854)

QY 192 LysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeu 211
Db 1798 AAGCAGATGAAGCTTCTACGAAACACAGCAGCTGTGTGCAAGAGAGACCTG 1739
QY 212 ArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGlu 231
Db 1738 CGCGGTGAGCAGCAGCAGCC-GTCTGCGCCGAGCAAGCTTGAGAGCCTATGCGGTGAG 1680
QY 232 LeuGlnArgHisAsnArgSerLeuLysGluGluValGlnArgAlaArgGluGluGlu 251
Db 1679 CTGACGGGCAACACCGCTCCCTCAAGAGAGAGGTGTGAGCGGGCCCGAGGAGGAG 1620
QY 252 GluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGln 271
Db 1619 GAGAGCGCAGGAGGTGACCTCCACTTCCAGGTGACACTGAATGACATTCAGCTGCAG 1560
QY 272 MetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGlnMetGluLeuAlaGlu 291
Db 1559 ATGGAACAGCACAATGACGCGCACTCCAAAGCTGCGCCAGAGAGACATGGAGCTGGGTGAG 1500
QY 292 ArgLeuLysLysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAspLysValPhe 311
Db 1499 AGGCTCAAGAGCTGATTTGACGATATGAGCTGCGCAGGAGCATATCGACAAGCTTC 1440
QY 312 LysHisLysAspLeuGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMet 331
Db 1439 AACACAAGACCTTACACAGCAGCTGTGTGATGCCAAGCTCCAGCAGGCGCCAGGAGATG 1380
QY 332 LeuLysGluAlaGluGluArgHisGlnArgGluLysAspPheLeuLysGluAlaVal 351
Db 1379 CTAAGGAGGAGAGAGGCGCACCGCGGAGAGGATTTTCTCTGAAGAGGAGGTA 1320
QY 352 GluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLysLysGlnLeu 371
Db 1319 GAGTCCAGCAGGATGTGTGAGCTCATGAAGCAGCAGACCCACCTTGAAGCAACAGCTT 1260
QY 372 AlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluVal 391
Db 1259 GCCCTATACACAGAAAGTTTGAGGRTTCCAGAACACACTTTCCAAAAGCAGCAGGTA 1200
QY 392 PheThrThrPheLysGlnGluMetGluLysMetThrLysLysIleLysLysLeuGluLys 411
Db 1199 TTCACCATTCAGCAGGAGATCGAAAGATGACTAAGAGATCAAGAGCTGGAGAA 1140
QY 412 GluThrThrMetTyrArgSerArgTyrGluSerSerHisLysAlaLeuLeuGluMetAla 431
Db 1139 GAAACCAACCATGTACCGGTCCGCGTGAGAGCAGCAACAGGCCCTTGTGTGATGGCT 1080
QY 432 GluGluLysThrValArgAspLysGluLeuGluGlyLeuGlnValLysIleGlnArgLeu 451
Db 1079 GAGAGAAACAGTCGCGGATTAAGAACTGGAGGCTTCAGGTAAATCCAAAGCTG 1020
QY 452 GluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgValGlnAsp 471
Db 1019 GAGAAGCTGTCCGGGCACTCAGACAGAGCGCAATGACCTGAACAGAGGGTACAGGAC 960
QY 472 LeuSerAlaGlyGlnGlnGlySerLeuThrAspSerGlyProGluArgArgProGlu-Gl 491
Db 959 CTGAGTGTGTGGCCAGGGCTCCCTCACTGACAGTGGCCCTGAGAGAGCCGAGAGGG 900
QY 491 yProGlyAla-GlnAlaProSerProArgValThrGluAlaProCysTyrProGlyA 511
Db 899 CACTGGGGCATCAAGCACCAGCTCCCGGAGTCCAGAGAGCGCTTGTGTCCAGGAG 840
QY 511 la-ProSerThrGluAlaSerGlyGlnThrClyProGlnGluProThrSerAlaArgAla 530

Db 839 CAACCGCAGCAGCAGCATCATCAGCCAGCTGG-CCTCAAGAGCCCACTCGCCAGGGCC 781

RESULT 9
US-08-968-751-5
; Sequence 5, Application US/08968751
; Patent No. 5948643
; GENERAL INFORMATION:
; APPLICANT: Rubinfield, Bonnie
; APPLICANT: Polakis, Paul G.
; APPLICANT: Ligensfelter, Carol
; APPLICANT: Vuong, Terilyn T.
; TITLE OF INVENTION: MODULATORS OF BRCA1 ACTIVITY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ONYX Pharmaceuticals, Inc.
; STREET: 3031 Research Drive
; CITY: Richmond
; STATE: CA
; COUNTRY: USA
; ZIP: 94806
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/968,751
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Giotta, Gregory
; REGISTRATION NUMBER: 32,028
; REFERENCE/DOCKET NUMBER: ONYX1024 GG
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 262-8710
; TELEFAX: (510) 222-9758
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1191 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 34..1191
US-08-968-751-5

Alignment Scores:
Pred. No.: 4,7e-83 Length: 1191
Score: 1047.50 Matches: 230
Percent Similarity: 69.01% Conservative: 55
Best Local Similarity: 55.69% Mismatches: 104
Query Match: 38.77% Indels: 24
DB: Gaps: 6

US-10-023-523-8 (1-530) x US-08-968-751-5 (1-1191)

QY 3 SerProGlyGlnProGluAlaGlyProGlu-GlyAlaGlnGluArgProSerGlnAlaA 22
Db 12 TCCCCCGGGCTGCAGAAATTCGGCAGGAGGCGGCCGAGAG----- 54
QY 22 aProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLys-----ProGl 40
Db 55 ----CGCACTGAGCGCGGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 110
QY 40 uClyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSe 60
Db 111 TGGCAATATGGAAGAGCTGGAAATTTGTGGCTAGGGGTGAACAGCATATGTTGTGTAA 170
QY 60 rArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyProGl 80
Db 171 CTCTCAATCAATGATATTTCTTCAACATCAA-----GGCTCAATTTGTTGGTGCACAA 224

Db	16792	CCTGAGAGGAGCCAGAGGGGCTGGGGCTCAAGCACCAGCTCCCCCAGGGTCACAGAA	16851
Qy	505	AlaProCysThrProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGlu	524
Db	16852	GCCCCTGTGTACCCAGGACCCGAGCACAGAAGCATCAGGCCACTGGGGCTCAAGAG	16911
Qy	525	ProThrSerAlaArgAla	530
Db	16912	CCCACCTCGGCCAGGGCC	16929

RESULT 11

```

US:09-833-381-1333
; Sequence 1333, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nuclieic
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 1333
; LENGTH: 614
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(614)
; OTHER INFORMATION: n = A,T,C or G
US:09-833-381-1333

```

Alignment Scores:	1.02e-48	Length:	614
Pred. No.:	62.50	Matches:	144
Score:	83.66%	Conservative:	25
Percent Similarity:	71.29%	Mismatches:	25
Best Local Similarity:	24.15%	Indels:	8
Query Match:	4	Gaps:	1
DB:			

US-10-023-523-8 (1-530) x US-09-833-381-1333 (1-614)

150	QY	LysGlyLeuGlyGlyGluLeuThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrPro	169
9	Db	AAACCTTTAGGAAAAAGAACTTTATTTACTGATGCAAGCCCTAAACACCCCTTTCAAACCCCA	68
170	QY	GluGluLysLeuAlaLeuCysIstysIstysTrAlaGluLeuLeuGluGluHisArgAsn	189
69	Db	CACCACAANCTGCGAGCTCTCTGTAGAAATATGCTGATCTTCTGGAGGAGACGAGAGT	128
190	QY	SerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAsp	209
129	Db	GTTTCAGAGCAAAATGAAGATCTCGCAGAAAGCAAGCCAGCATTTGTGAAAGAGAAAGTT	188
210	QY	HisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeu-Cy	229
189	Db	CACCTTGCAGAGTGAACATAGCAAGGCTACTTGCACAAGCAAGCTAGAAATCTCTTTTG	248
229	QY	sArgGluLeuGlnArgHisAsn--ArgSerLeuLysGluGluGlyValGlnArgAlaArg	248
249	Db	CAGAGAACTTCAGCGTCCACATTAAGCCGTTAAAGAGGAAATATGCGCANGCACNA	308
249	QY	GluGluGluGlu-LysArgLysGluValThrSerHisPheGlnValThrLeuAsnApril	268
309	Db	GAGGAAGAAGAACCCGACGTAAAGAAGCAACTGTCACATTTCCAGATTACCTTAAATGAAAT	368
268	QY	ecGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGln	288
369	Db	TCAAAGCCCGAGCTGGAGCAGCATGATCCCAACAGCCCAACTCCGACAGAGAAACATTGA	428
288	QY	uLeuAlaGluArgLeuLysLysLeuLeuGluGlnTrpGluLeuArgGluGluHisIleAs	308

[illegible]

RESULT 12

US-08-633-3068A-5
; Sequence 5, Application US/08533306A
; Patent No. 5837457
; GENERAL INFORMATION:
; APPLICANT: Liu, Pu
; APPLICANT: Collins, Francis S.
; APPLICANT: Siciliano, Michael J.
; APPLICANT: Claxton, David
; TITLE OF INVENTION: Markers for Detection of Chromosome 16
; TITLE OF INVENTION: Rearrangements
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
; STREET: P.O. Box 828
; CITY: Bloomfield Hills
; STATE: MI
; COUNTRY: USA
; ZIP: 48303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/533.306A
; FILING DATE: September 25, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Deann F.
; REGISTRATION NUMBER: 36683
; REFERENCE/DOCKET NUMBER: 21115-00869COB

```

TELECOMMUNICATION INFORMATION:
TELEPHONE: (810) 641-1600
TELEFAX: (810) 641-0270
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2680 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
INDIVIDUAL ISOLATE: Sample 2
TISSUE TYPE: Acute myelomonocytic leukemia, M4EO
TISSUE TYPE: subtype (inv16)
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 16(inv16)(p13q22)]
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2451
US-08-533-306A-5

Alignment Scores:
Pred. No.: 7,29e-15
Score: 278.50
Percent Similarity: 38.90%
Best Local Similarity: 22.99%
Query Match: 10.31%
Length: 2680
Matches: 140
Conservative: 10%
Mismatch: 22%
Indels: 16%

```

DB: 2 Gaps: 25

US-10-023-523-8 (1-530) x US-08-533-306A-5 (1-2680)

Qy 3 SerProGlyGlnPro-----GluAlaGlyProGluGlyAlaGlnGluArgPro 18
Db : : : : :
582 GTCCAGGACACCAGGAGTTGCTTCAAGAAGAAACCGGCAGAAGCTCAA---CGTGTC 638
Qy 19 SerGlnAlaAlaProLavalGluAlaGluGlyProGlySerSerGlnAlaProArgLys 38
Db : : : : :
639 TACGAAGCTCGGCCAGCT-----GGAGGAGGCGGAACAGCCTCGCAAGA 683
Qy 39 ProGluGly---AlaGlnAlaArgThrAlaGlnSerClvAlaLeuArgAspValSerGlu 57
Db : : : : :
684 CCAGCTGGACGAGAGATGAGGCGCAAGCAGAACCTGGAGCGGCACAATCTCCACTCTCAA 743
Qy 58 GluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrcysValaAspAsnAsnGlnGly 77
Db : : : : :
744 CATCCAGCTCTCGACTCGAAGAGAGCTGCAGGACTTTGCCAGCACCGGTGAAGCTCT 803
Qy 78 GlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluPro----- 91
Db : : : : :
804 GGAAGAGGGGAAGAAGAGGTTCCCAAGAGAGATCGAGAACCTCACCAGCAGTACGAGGA 863
Qy 92 -----GluAspAlaGluLysSerArgThrTyrcysVala 102
Db : : : : :
864 GAAGCGCGCGCTTATCATTAACATGGNAAGACCAAGAACAGGCTTCAGCAGGAGCTGGA 923
Qy 103 ArgAsn-----GlyGluProGluProThrProValValTyrglyGluLysGlu 118
Db : : : : :
924 CGACCTGGTTGTTGATTGTGACACACCAAGCGGCAACTCGTCTCCAACCTGGAAGAAGCA 983
Qy 119 Pro-----SerLyseGlyAspProAsnThrGluGluIleArgGlnSer 132
Db : : : : :
984 GAGGAATTGATCAGTTGTTAGCCGAGGAGAAAAACATCTCTTCCAAATACGC----- 1037
Qy 133 AspGluValGly-AspArgAspHisArgArgProGlnGluLysLysLysAlaLysGlyLe 152
Db : : : : :
1038 -----GGATGAGAGGCACAGAGCTGAGGCAGAAAGCCAGGAGGAAG----- 1077
Qy 152 uGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLy 172
Db : : : : :
1078 -----GAAACCAA 1085
Qy 172 sLeuAlaAlaLeuCysLysLysTyrrAlaGluLeuLeuGluHisaArgAsnSerGlnLy 192
Db : : : : :
1086 GGCCCTGCTCCCTGGCTCGGCCCTTGAAGAGGCCCTTGAAGCCAAAGAGAACTCGAGCG 1145
Qy 192 sGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuAr 212
Db : : : : :
1146 GACCAACAAAATGCTCAAAACCGAAATGGAAGACCTCGTCAAGTCCAAAGATCAGCTGGG 1205
Qy 212 gGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeucysArgGluLe 232
Db : : : : :
1206 CAAGAACGTCATGAGCTGGAGAGTCCAGCGGCCCTTGAGACCCCAGATGAGGAGAT 1265
Qy 232 uGlnArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAlaargGluGluGl 252
Db : : : : :
1266 GAAGACGCGAGCTGGAAGAGCTGGAGGACGAG-----CTCAAGACCTCGGAGGACGCCAA 1319
Qy 252 uLysArgLysGluValThr-----SerHisPheGlnValThrLeuAsnAs 267
Db : : : : :
1320 ACTCGCGCTGGAAGTCAACATGCAAGCGCTCAAGCGCCAGTTCCAA-----AGGGA 1370
Qy 267 pIleGlnLeuGlnMetGlnGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMe 287
Db : : : : :
1371 TCTCCAAGCCCGGACGAGCAGAGATGAGGAGAGAGGAGGCAACTGTCAGACAGACAGCTTCA 1430
Qy 287 tGlu-----LeuAlaGluArgLeuLy 294
Db : : : : :
1431 CGAGTATGAGACGGAACTGGNAGACGAGCGGAACGAACGTGGCCCTGCAGCTGCAGCAAA 1490
Qy 294 sLysLeuIleGlu-----GlnTyrgluLeu----- 302

RESULT 13
US-08-742-923A-5
Sequence 5, Application US/08742923A
Patent No. 5869611
GENERAL INFORMATION:
APPLICANT: Liu, Pu
APPLICANT: Collins, Francis S.
APPLICANT: Siciliano, Michael J.
APPLICANT: Claxton, David
TITLE OF INVENTION: Markers for Detection of Chromosome 16
TITLE OF INVENTION: Rearrangements
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Harness, Dickey & Pierce, P.L.C.

STREET: P.O. Box 828
 CITY: Bloomfield Hills
 STATE: MI
 COUNTRY: USA
 ZIP: 48303
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/742,923A
 FILING DATE: No. 5869611 member 1, 1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Smith, Deann F.
 REGISTRATION NUMBER: 36683
 REFERENCE/DOCKET NUMBER: 2115-00869DVC
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (810) 641-1600
 TELEFAX: (810) 641-0270
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2680 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 INDIVIDUAL ISOLATE: Sample 2
 TISSUE TYPE: Acute myelomonocytic leukemia, M4E0
 TISSUE TYPE: subtype (inv16)
 POSITION IN GENOME:
 CHROMOSOME/SEGMENT: 16 [inv(16) (p13q22)]
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..2451
 US-08-742-923A-5

Alignment Scores:
 Pred. No.: 7,29e-15 Length: 2680
 Score: 278.50 Matches: 146
 Percent Similarity: 38.90% Conservative: 101
 Best Local Similarity: 22.99% Mismatches: 225
 Query Match: 10.31% Indels: 164
 DB: 2 Gaps: 25

US-10-023-523-8 (1-530) x US-08-742-923A-5 (1-2680)
 QY 3 SerProGlyGlnPro-----GluAlaGlyProGluGluAlaGlnGluArgPro 18
 Db 582 GCTCCAGGACCCAGGAGTTGTTCAAGNAGNAAACCGCGCAGAACTCAA---CGTGTC 638
 QY 19 SerGlnAlaProLaValGluAlaGluGlyProGlySerSerGlnAlaProArgLys 38
 Db 639 TACGAAGCTCGCCAGCT-----GGAGGAGGAGCGGAGCAGCTGCAAGA 683
 QY 39 ProGluGly---AlaGlnAlaArgThrAlaGlnSerGlyValaLeuArgAspValSerGlu 57
 Db 684 CCAGCTGGACGAGAGATGAGCGCCAGCAGCAACCTGGAGCGCCACATCTCCACTCTCAA 743
 QY 58 GluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrcysValAspAsnGlnGly 77
 Db 744 CATCCAGCTCTCCGACTCGAAGAGAGAGCTGCAGGACTTTGCCAGCACCGTGAAGCTCT 803
 QY 78 GlyProGlyAspGlyAlaGlnGlyGluProAlaGluPro----- 91
 Db 804 GGAAGAGGGGAGAGAGGTTCCAGAGAGGAGATCGAGAACCTCACCAGCAGTACGAGGA 863
 QY 92 -----GluAspAlaGluLysSerArgThrTyrrValaLa 102

Db 864 GAAGCGCGCGCTTATGATAAACTGGAAAAGACCAAGAACAGGCTTCAGCAGAGCTGCA 923
 QY 103 ArgAsn-----GlyGluProGluProThrProValValTyrrGlyGluLysGlu 118
 Db 924 CGACCTGGTTGTTGATTTGGACCAACGAGCGGCAACTCGTGTCCAACCTGGAAGAAGCA 983
 QY 119 Pro-----SerLysGlyAspProAsnThrGluGluLeuArgGlnSer 132
 Db 984 GAGGAATTGTGATCAGTTGTAGCCGAGGAGAAAACATCTCTTCCAAATACGC----- 1037
 QY 133 AspGluValGly-AspArgAspHisArgArgProGlnGlnGluLysLysLysLysGlyLe 152
 Db 1038 -----GGATCAGAGGACAGAGCTGAGGAGAGAGCCAGGAGAG----- 1077
 QY 152 uGlyLysGluLeuThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLys 172
 Db 1078 -----GAAACCAA 1085
 QY 172 sLeuAlaLeuCysLysLysTyrrAlaGluLeuLeuGluGluHisArgAsnSerGlnLys 192
 Db 1086 GGCCTGTCTCTGCTCGGCCCTTGAAGAGGCTTGGAAAGCCAAAGAGAGAACTCGAGCG 1145
 QY 192 sGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuAr 212
 Db 1146 GACCACAAATGCTCAAGCCGAAATGGAAGACCTGTCAGCTCCAGGATACGTGGG 1205
 QY 212 gGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLysCysArgGluLe 232
 Db 1206 CAAGAAGCTCCATGAGCTGGAGAGTCCAAAGCGGCCCTGGAGACCAGATGAGGAGAT 1265
 QY 232 uGlnArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGluGluGlu 252
 Db 1266 GAAGACGACGCTGGAAGAGCTGGAGAGCAG-----CTCAAGCTCGGAGGAGCGCAA 1319
 QY 252 uLysArgLysGluValThr-----SerHisPheGlnValThrLeuAsnAs 267
 Db 1320 ACTCGGCTGGAAGTCAACATGACGAGCGCTCAAGGCGCAGTTGAA-----AGGA 1370
 QY 267 pIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMe 287
 Db 1371 TCTCAAGCCCGGACGACGAGATGAGAGAGAGGAGGAGCACTGCAGAGACAGCTTCA 1430
 QY 287 tGlu-----LeuAlaGluArgLeuLys 294
 Db 1431 CGAGTATGAGACGGAACCTGGAGAGACGAGCAACAGAGCTGCCCTGGCAGCTGCAGCAA 1490
 QY 294 sLysLeuIleGlu-----GlnTyrrGluLeu----- 302
 Db 1491 GAAGAGCTGGAAGGGGACCTGAAAGACCTGAGCTTCAGGCCGACTCTGCCATCAAGGG 1550
 QY 303 -ArgGluHisIleAspLysValPheLys-----HisLysAspLeuGlnG 318
 Db 1551 GAGGAGGAGGAGCCATCAAGCAGCTACGCAAACTCAGGCTCAGATGAGGAGCTTCAAG 1610
 QY 318 nGlnLeuValAspAlaLys----- 324
 Db 1611 AGAGCTGGAAGATGCCCTGCTCCAGAGATGAGATCTTTGCCACAGCAAGAGAAATGA 1670
 QY 325 -----LeuGlnGlnAlaGlnLeuMetLeuLysGluAlaValG 335
 Db 1671 GAAGAAAGCCAAAGACTTGGAAAGCAGACCTCATCAGCTACAGAGGAGCTTCGCCGCCG 1730
 QY 335 sGluLysArgHisGlnArg-----GluLysAspPheLeuLeuLysGluAlaValG 352
 Db 1731 TGAGAGGCTGCCAAACCAAGCGGACCTCGAGAGAGAACTGGCAGAGGAGCTGGCCAG 1790
 QY 352 uSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLysLysGlnGlnLeuAl 372
 Db 1791 TAGCTGTCTGGAGAGAAACGACCTCCAGCAGAGAGAGCGCCCTGGAGGCGCCGATCGC 1850
 QY 372 aLeuTyrrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluValP 392

Db 1353 GACCAACAAAATGCTCAAAGCCGAATGGAAGACCTGCTCAGCTCCAGGATGACGTGGG 1412
Qy 212 gGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLe 232
Db 1413 CAAGAACCTCCAGCTGAGAGAGTCAACGGCGCTGAGAGCCAGATGGAGGAGAT 1472
Qy 232 uGlnArgHisAsnArgSerLeuLysGluGluValGlnArgAlaArgGluGluGlu 252
Db 1473 GAAGACGCGAGTGAAGAGCTGGAGGACGAG-----CTGCAAGCTCGGAGGACGCCAA 1526
Qy 252 uLysArgLysGluValThr-----SerHisPheGlnValThrLeuAsnAs 267
Db 1527 ACTCGGCTGAGAGTCAACATGCGAGCGCTCAAGGGCCAGTTCGAA-----AGGA 1577
Qy 267 pIleGlnLeuGlnMetGluGlnHisAsnGluArgSerLysLeuArgGlnGluAsnMe 287
Db 1578 TCTCCAGCCGGGACGAGCAGAGATGAGGAGAGGAGGCACTGCAGAGACAGCTTCA 1637
Qy 287 tGlu-----LeuAlaGluArgLeuLys 294
Db 1638 CGAGTATGACGAACTGGAAGACGAGCGGAAACGAACTGCTCGCTGGCAGTGCAGCAA 1697
Qy 294 sLysLeuIleGlu-----GlnTyrGluLeu----- 302
Db 1698 GAAGAGCTGGAAGGACCTGAAGACCTGAGCTTCAGGCGCACTCTCCATCAGGG 1757
Qy 303 -ArgGluGluHisIleAspLysValPheLys-----HisLysAspLeuGlnG 318
Db 1758 GAGGAGGAGAACCATCAGCAGCTACGCAAACTCGAGGCTCAGATGAAGGACTTCAAAG 1817
Qy 318 nGlnLeuValAspAlaLys----- 324
Db 1818 AGAGCTGAAGATGCCGCTGCTCAGAGATGAGATCTTGGCACACCCAAAGAGAATGA 1877
Qy 325 -----LeuGlnGlnAlaGlnGluMetLeuLysGluAl 335
Db 1878 GAAGAAAGCCAAAGCTTGAAGCAGACCTCATCAGCTACAAGAGGACCTCGCGCGCG 1937
Qy 335 aGluGluArgHisGlnArg-----GluLysAspPheLeuLysGluAlaValG 352
Db 1938 TGAGAGGCTCGCAAAAGCGGACCTCGAGAGGAGGAACTGCGAGAGAGCTGCGCAG 1997
Qy 352 uSerGlnArgMetCysGluLeuMetLysGlnGlnGlnThrHisLysLysGlnGlnLeuAl 372
Db 1998 TAGCTCTCGGAGAGAACCACTCCAGAGCAGAGAGCGCGCTGGAGCCCGATCGC 2057
Qy 372 aLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluValPh 392
Db 2058 CCAGCTGGAGAGAGCTGGAGGAGGAGCAGGCGCAACATGAGGCGCCATGAGCGAC----- 2112
Qy 392 eThrThrPheLysGlnGluMetGluLysMetThrLysLysIleLysLysLeuGluysG 412
Db 2113 -----CGGTCCGCAAGGCCACACAGCGCGCAGAGCTCAGCAACGA 2156
Qy 412 uThrThrMetTyrArgSerArgTrpGluSerSerAsnLysAlaLeuLeuGluMet---Al 431
Db 2157 GCTGGCCACAGAGCGCCAGCAGCGCCCAAGAAATGAGATGAGTGCCTGGCAGAGCTCGAGCG 2216
Qy 431 aGluGluLysThrValArgAspLys-----GluLeuGluGly----- 443
Db 2217 GCAGAACAGAGAGCTCCGGAGCAGAGTCCACAGATGAGGGGGCGGCTCAAGTCCAAGTT 2276
Qy 444 -----LeuGlnValLysIleGlnArgLeuGluLysLeu----- 454
Db 2277 CAAGTCCACCATCGCGCGCTGGAGGCCAAGATTGCACAGCTGGAGGAGCGGTGCGACA 2336
Qy 455 -----CysArgAlaLeuGlnThrGluArgAsnAspLeu-----AsnLysArgValG 470
Db 2337 GGAGGCCAGAGAGAAACAGCAGCCACCAAGTCCGTGAGCAGAGAAAGACAGAGTGA 2396
Qy 470 nAspLeuSerAlaGlyGlnGlnGlySerLeuThrAsp-----SerG 484
Db 2397 GGAATCTT-GCTGCGAGTGGAGACGAGCGCAAGATGGCCGAGCAGTACAGGAGCAGG 2455

Qy 484 yProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrCl 504
Db 2456 CAGAGAAAGCATGTCAGGCTCAACAGCTCAAGAGGAGCTGGAGGAGGAGGAGGAGG 2515
Qy 504 uAlaProCysTyrProGlyAlaProSerThrGluAlaSerGly 518
Db 2516 AGTCCAGCGCATCAACAGCCCAACCGCAGGAGCTGCAGCGGGA 2558
RESULT 15
US-08-742-923A-3
; Sequence 3, Application US/08742923A
; Patent No. 5869611
; GENERAL INFORMATION:
; APPLICANT: Liu, Pu
; APPLICANT: Collins, Francis S.
; APPLICANT: Siciliano, Michael J.
; APPLICANT: Claxton, David
; TITLE OF INVENTION: Markers for Detection of Chromosome 16
; TITLE OF INVENTION: Rearrangements
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
; STREET: P.O. Box 828
; CITY: Bloomfield Hills
; STATE: MI
; COUNTRY: USA
; ZIP: 48303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,923A
; FILING DATE: No. 5869611member 1, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Dean F.
; REGISTRATION NUMBER: 36693
; REFERENCE/DOCKET NUMBER: 2115-00869DVC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 641-1600
; TELEFAX: (810) 641-0270
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2887 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; INDIVIDUAL ISOLATE: Sample 1
; TISSUE TYPE: Acute myelomonocytic leukemia, M430
; TISSUE TYPE: subtype (inv16)
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: 16[inv(16) (p13q22)]
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2658
US-08-742-923A-3

Alignment Scores:
Pred. No.: 8,146-15
Score: 278.50
Percent Similarity: 38.90%
Best Local Similarity: 22.9%
Query Match: 10.31%
DB: 2
Length: 2887
Matches: 146
Conservative: 101
Mismatches: 225
Indels: 164
Gaps: 25

CITY: Newport Beach
 STATE: CA
 COUNTRY: U.S.A.
 ZIP: 92660
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/056,200
 FILING DATE: 30-APR-1993
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Pedrick, Michael P.
 REGISTRATION NUMBER: 36,799
 REFERENCE/DOCKET NUMBER: NIH054.001A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (714) 760-0404
 TELEFAX: (714) 760-9502

INFORMATION FOR SEQ ID NO: 93:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 9551 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1507..1644
 NAME/KEY: intron
 LOCATION: 1645..2511
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 2512..8070

US-08-056-200-93

Alignment Scores:
 Pred. No.: 1,19e-13 Length: 9551
 Score: 274.00 Matches: 140
 Percent Similarity: 40.73% Conservative: 95
 Best Local Similarity: 24.26% Mismatches: 235
 Query Match: 10.14% Indels: 109
 DB: 1 Gaps: 20

US-10-023-523-8 (1-530) x US-08-056-200-93 (1-9551)

QY 5 GlyGlnProGluAlaGlyProGluGly-----AlaGlnGluArg---ProSerGln 20
 Db 3021 GGAGCTGAGGAGGAAGGCGCGAGGAGAAACAGCAGCAAGAGCGGAGAGCGCCCAAGACAG 3080
 QY 21 AlaAlaProAla----- 24
 Db 3081 AGTGTTCAGGAGGAGGAAGAAAGAGCTGAGGAGGAGCGCGAGACAGTGTCTCCGAAGGA 3140
 QY 25 ---ValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGln 43
 Db 3141 AGAAGAGAAGTTGCAGGAAGAGGAGCGCGCAGCGGCAAGAGAGAGCTCCAGGAGGAAGAAGA 3200
 QY 44 AlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeu 63
 Db 3201 GCAGTACGGAAGCTGGAGCGCGCAAGAGCTGAGGAGGAGGAGCGCCAGGAGGAAGAGCAGCA 3260
 QY 64 GluAspIleLeuSerThrTy-CysValAspAsnAsnGlnGlyProGlyGluAspGly 83
 Db 3261 GCAGCAAGGCTGAGCGCGAGCAGCAACTAAGCGCAAGCAGGAGGAGGAGCGCGGA 3320
 QY 84 AlaGlnGly-GluProAlaGluProGluAsp-----AlaGluLysSerAr 98
 Db 3321 GCAGCAGGAGGAGGCGCGAGCAGCAGGAGGAGGCGCGAGCAGCAGGAGGAGGCGCGCA 3380

QY 98 gThrTyrValAlaAlaArgAsnGlyGluProGluProThrProValValTyrGlyGluLysGln 118
 Db 3381 GCAGCAGCTGAGCGCGAGCAGCAGGAGAGGCGCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 3440
 QY 118 uProSerLysGlyAspProAsnThrGluGluLeuArg-----GlnSerAspGluValGln 136
 Db 3441 GGAGGAGAGCGCGAG-----CAGCAGCTGAGCGCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 3491
 QY 136 yAspArgAspHisArgArgProGlnGlnLysLysLysAlaLysGlyLeuGlyLysGluLys 156
 Db 3492 CGAGCAGCAGCTGAGCGCGAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3534
 QY 156 eThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaAlaLeu 176
 Db 3535 -----CAGCAGCTGAGCGCGA-CAGCAGCTGAGCGCGCGAGCAGCAGCAGCAGCAGCAGCAG 3580
 QY 176 uCysLysLysTyrAlaGluLeu-LeuGluHisArgAsnSerGlnLysGlnMetLysLys 196
 Db 3581 GCGCGCAGCAGCAGCTGAGCGCGAGCAGCAGCTGAGCGCGAGCAGCAGCAGCAGCAGCAGCAGCAG 3640
 QY 196 eLeuGlnLysLysGlnSerGlnLeuValGlnGlnLysAspHisLeuArgGlyGluHis 216
 Db 3641 AGCAGCAGCTGAGCGCGAGCAGCAGCTGAGCGCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 3700
 QY 216 erLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisA 236
 Db 3701 AG-----CTGAGCGCGAGCAGCAGGAGGAG-----AGGCAGCAGCAGCAGCAGCAGCAG 3745
 QY 236 snArgSerLeuLysGluGlyValGlnArgAlaArgGluGluGluGlu-----L 253
 Db 3746 AGCAGGAGGCGCGAGCAGCGCTGAAGCGCGAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3805
 QY 253 ysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetG 273
 Db 3806 AGCGCGAGGAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 3865
 QY 273 luGlnHisAsnGluArgAsnSerLysLeu-----ArgGlnGluAsnMetGluL 289
 Db 3866 ACCAGGAGGAGGAGGAGCGCGAAGCTGTGCTGAAGCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3925
 QY 289 euAlaGluArgLeuLysLysLeuLeu-----GluGlnTyrGluLeuArgGluGluHisI 307
 Db 3926 AGCAGGAGGCGCGAGCAGCAGCACTAAGCGCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 3985
 QY 307 leAspLysValPheLysHisLysAspLeuGlnGlnGlnLeu----- 320
 Db 3986 TGAAGCGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4045
 QY 321 -----ValAspAlaLysLeuGlnGlnAlaGlnGlnMetLeuLysG 334
 Db 4046 GAGCGCAGCAGGAGGAGGAGCTCGAGCAGCTGTGAAGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAG 4105
 QY 334 luAlaGluGluArgHisGlnArg-----GluLysAspPheLeu 347
 Db 4106 AGCAGGAGGCGGAGCAGCAGCGCTGAAGCGCGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4165
 QY 347 euLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlnThrHisL 367
 Db 4166 TGAAGCGCGAG---GAGGAGGCGCGCAGCAGCGCTGAAGCGCGAGCAGGAGGAGGAGGAGGAGGAG 4222
 QY 367 euLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerL 387
 Db 4223 TCAGCAGCGAGCTGAG-----CGCGAGGAGGTGAGAGAGCTCGAGCAGGAG---GAGA 4273
 QY 387 ysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysIleL 407
 Db 4274 GCGCGCAGCAGCGCTGAAGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4333
 QY 407 ysLysLeuGluLysGluThrMetTyrArgSerArgTrpGluSerSerAsnLysAlaL 427
 Db 4334 AGAGCGGAGGAGCAGGAG-----GAGAGCGCGCAGCAGCAGCACTGAGCGCGAGCAGCAG 4381

RESULT 21
US-09-480-884A-119
; Sequence 119, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.

QY 256 ---GluValThrSerHisPheGlnValThr----- 264
 Db 5579 CTTGGAACCTAGGAGCGAGCTGCAGATCAGCAACACCGGACCTTGAAGTGCAGGCGGT 5638
 QY 265 -LeuAsnAspIleGlnLeuMetGluGlnHisAsnGluArgAsnSerLysLeuArgG1 284
 Db 5639 GATTATGATTACAGAGAGAGAGGGAATTCAGACAGGAATTCAGAAATTCACAAA 5698
 QY 284 nGluAsnMetGluLeuAlaGluArgLeuLysLeuLeuGln---TyrGluLeuAr 303
 Db 5699 GCAGGCTTTAGAGCATCTATAGATTAGGATTCAGGAATCAAGAATCAGTGTACTCAGGTGGT 5758
 QY 303 gGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeuValAspAl 323
 Db 5759 ACAGGAAGAGAGAGCGCTTCGTGGAATCAAGCTCTGGAGCAAGAC-----AAGGC 5812
 QY 323 aLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluArgHisGlnArgGluLyl 343
 Db 5813 AAGCTGCAGAGGCTGCAGGATGAGCTGAATCGTGA----- 5849
 QY 343 sAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnG1 363
 Db 5850 -----AAATCAACTCTAGAGGCA----- 5867
 QY 363 nGluThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAs 383
 Db 5868 -GAACACAGGCTGAACAGCGCTGGAG-----TGTGAGAA---CAGCAATTCAGAA 5917
 QY 383 nThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetTh 403
 Db 5918 TGACCTGAATCAG-----TGGAAGACTCAATATATCCCGCAAGGA 5956
 QY 403 rLysLysIleLysLysLeuGluLysGluThrThrMetTyrArgSerArgTyrGluSerSe 423
 Db 5957 GGAGGCTATTAGGAATAGATTCGGAA----- 5984
 QY 423 rAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLys----- 439
 Db 5985 -----AGAGAAAAGAGTGCAGAGAGAGAGAAACAGCTCTTAG 6019
 QY 440 ---GluLeuGluGlyLeuGlnValLysIleGlnArgLeuGluLysLeuCysArg----- 456
 Db 6020 GAGTGAGATCGAAGACTCCACAGAGATCAAGAGATTGAAGAGAGGTGCAGCGCTAA 6079
 QY 457 -----AlaLeuGlnThrGluArgAsnAspLeuAs 466
 Db 6080 GCTGGAGGATTCACCGAGGAGACAGTCACAGTTAGAAACAGAACGCTCCCGATATCA 6139
 QY 466 nLysArgValGlnAspLeuSerAlaGlyGlnGlySerLeuThrAspSerGlyProG1 486
 Db 6140 GAGGGAGATTGATAACT-----AG 6160
 QY 486 uArgArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGluAlaPr 506
 Db 6161 ACAGCGCCCATATGG-----TCCCATCGAGAGACCCAGACTGA 6199
 QY 506 oCys 507
 Db 6200 GTGT 6203

RESULT 25

US-08-875-435B-5
 ; Sequence 5, Application US/08875435B
 ; Patent No. 6593304
 ; GENERAL INFORMATION:
 ; APPLICANT: Hasegawa, Kazuhide
 ; APPLICANT: Arakawa, Emi
 ; APPLICANT: Oda, Shoji
 ; APPLICANT: Matsuda, Yuzuru
 ; APPLICANT: Takahashi, Katsuhito
 ; APPLICANT: Suganara, Michihiro
 ; APPLICANT: Ishiyama, Haruo

; TITLE OF INVENTION: RECOMBINANT DNA COMPRISING DNA CODING
 ; TITLE OF INVENTION: FOR MYOSIN HEAVY CHAIN SM1 ISOFORM PROTEIN INSERTED INTO
 ; TITLE OF INVENTION: VECTOR DNA, MICROORGANISM CARRYING THE RECOMBINANT DNA, AND
 ; TITLE OF INVENTION: AN AGENT FOR TREATMENT OF ARTERIOSCLEROSIS COMPRISING THE
 ; FILE REFERENCE: 07898-013001
 ; CURRENT APPLICATION NUMBER: US/08/875,435B
 ; PRIOR FILING DATE: 1997-07-25
 ; PRIOR APPLICATION NUMBER: PCT/JP96/00134
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 5
 ; LENGTH: 6644
 ; TYPE: DNA
 ; ORGANISM: Oryctolagus cuniculus
 ; US-08-875-435B-5

Alignment Scores:

Pred. No.: 8,72e-13 Length: 6644
 Score: 261.50 Matches: 157
 Percent Similarity: 37.54% Conservative: 96
 Best Local Similarity: 23.29% Mismatches: 216
 Query Match: 9.68% Indels: 208
 DB: 4 Gaps: 26

US-10-023-523-8 (1-530) x US-08-875-435B-5 (1-6644)

QY 9 AlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAla----- 22
 Db 3912 GCGGGCCGGCGGAGCTCAACGACAAAGTCCACAACTGCAGAATCAAGTGGAGAGCGT 3971
 QY 22 ----- 22
 Db 3972 CACGGCGATGCTCAGCAGGCCGAGGGAAGGCCATCAAGCTGGCCCAAGAGGTGGCGTC 4031
 QY 23 ProAlaValGlnAlaGluGly---ProGlySerSerGlnAlaProArgLysProGluGly 41
 Db 4032 CCTCGGTCCAGCTCCAGGATACCCAGAGAGCT-----GCTCCAGAAGAAACCCGGCA 4085
 QY 42 AlaGlnAlaArgThrAlaGlnSerGlyAla-LeuArgAspValSerGluGluLeuSerAr 61
 Db 4086 GAAGCTCAAGCTGTCCCAAGCTGGCGAGCTGGAGACGAGAGAACAGCCCTGCAGGA 4145
 QY 61 gGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGlyProGlyG1 81
 Db 4146 GCAGCTGCAGCAGGA----- 4160
 QY 81 uAspGlyAlaGlnGlyGluProAlaGluPro----- 91
 Db 4161 -GATGGAGGCAAGCAGAACCTGGAGCGGCACATCTCCACCTGGAACATCCAGCTCTCCG 4219
 QY 92 -----GluAspAlaGlu-----LysSerArgThrTy 100
 Db 4220 ACTCAAGAAGAAAGCTGCAGGACTTTGCCAGACACCTGGAGTCTCTGGAGGAAGCAAGA 4279
 QY 100 rValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGluLysGluProSe 120
 Db 4280 AGAGGTTCCAGAAGAAATTCAGAGCTCACCAGCAG--TACGAAGAGAAAGACAGCTGC 4337
 QY 120 r-----LysGlyAspProAsnThrGluGluIleArgGln-----SerAspGluValG1 136
 Db 4338 TTACGATAAACTGGAAAAGACCAGAACAGAGGCTTCAGCAGAGCTGGAGAGCTGCTGCTCT 4397
 QY 136 yAspArgAspHisArgArgProGlnGluLysLysLysAlaLysGlyLeuGlyLysLul1 156
 Db 4398 AGACTTGGATAACACGCGCAACTGTGTCCCACTCGGAAGAAAGACGACGAAGAGTTCGA 4457
 QY 156 eThrLeuLeuMetGlnThrLeuAsnThrLeuSer----- 167
 Db 4458 TCAGTTGTTAGCGGAGGAAAAGAACATCTCTTCCAAAGTATGCGGATGAAGGAGCCGAGC 4517
 QY 168 -----ThrProGluGluLysLeuAlaLeuCysLysLysTyAlaG1 182

Db 4518 CGAGCTGAAGCAAGGAAAGCAAGGCTTGTCTCTGCTGGCCCTCGAGGA 4577
Qy 182 uLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSe 202
Db 4578 GGCCTTGGAGCCCAAGAGAGCTCGAGAGAACCAACAAATGCTCAAGCCGAGATGGA 4637
Qy 202 rGlnLeuValGlnLysAspHisLysLeuArgGlyGluHisSerLysAlaValLeuAla 222
Db 4638 AGACCTCGTCAGCTCCAGAGACGACGTGGGCACAGACGTCATGATGAGAGAGTCCAA 4697
Qy 222 gSerLysLeuGluSerLysCysArgGluLeuGlnArgHisAsnArgSerLysLysGlu 242
Db 4698 GCGGCGCTTGGAGACACAGATGAGAGATGAAGACGACGCTGGAAGAGCTAGAGGACGA 4757
Qy 242 uGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThrSer 259
Db 4758 G-----CTGAGGCCACCGAGGACGCCAAGCTGCGGTGGAGGTCAACATGACGAGCCCT 4811
Qy 260 -----HisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnG 277
Db 4812 CAAGTCCAGTTCAG-----CGGATCTCCAGGCCCGGGATGATGAGCAGAGCA 4862
Qy 277 uArgAsnSerLysLeuArgGlnLeuAsnMetGlu----- 288
Db 4863 GAAGAGGAGCGAGCTGCAGAGCGAGCTGCATGATGACGACGAGCACTGGAAGACGAGCG 4922
Qy 289 -----LeuAlaGluArgLysLysLysLeuLeuLeu-----GlnTy 300
Db 4923 CAAGCAGCGGCCCTGGCCCGGCGAGCAAGAGAGAGCTGGAGGGGAGCTGAAAGACCT 4982
Qy 300 rGluLeu-----ArgGluGluHisLysLysValPhe 312
Db 4983 GGAGCTTCAGCGGACTCCGCCATCAAGGGCGGAGAGAGCCATCAAGCAGCTTCTGAA 5042
Qy 312 s-----HisLysAspLeuGlnGlnLeuValAspAlaLys----- 324
Db 5043 ACTGAGGCTCAGATGAAGACTTCCAGAGAGAACTGGAAGATGCCGCTCCAGAGA 5102
Qy 325 -----Le 325
Db 5103 CGAGATCTTCCACAGCCCAAGGAGACGAGAGCAAGCCAGAGTCTGGAGGACAGCT 5162
Qy 325 uGlnGlnAlaGlnLysMetLysLysGluAlaGluGluArgHisGlnArg-----G 342
Db 5163 CATGAGCTCAAGAGAGTCTGGCCGCGGAGAGAGAGGCTCGCAAAACAGGACAGCTTGA 5222
Qy 342 uLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGlnLeuMetLysG 362
Db 5223 GAAGAGAGCTGGCCGAGAGCTGGCCAGCAGCTTCTCGGAGAGAACCGCTGAGGA 5282
Qy 362 nGlnGluThrHisLysLysGlnGlnLeuAlaLeuTyThrGluLysPheGluPheG 382
Db 5283 TGAGAAGCGCGCTGGAGCGCGGATCGCACAGCTGGAGGAGGAGCTGGAGGAGGAACA 5342
Qy 382 nAsnThrLeuSerLys----- 387
Db 5343 GGCAACATGAGGCAATGACGACCGCTCGGACAGGCTACGACGAGCCGAGCAGCT 5402
Qy 388 -SerSerGluValPheThr-----ThrPheLysGlnG 398
Db 5403 CAGCAAGAGCTGCCACAGCGCCAGCACAGCCCAAGAAATGAGAGCCGAGCAGCA 5462
Qy 398 uMetGluLysMetThrLysLysLysLysLysLysLeuGluLysGluThrThrMetTy 417
Db 5463 GCTCAGCGGAGCAACCAAGAGCTCAAGAGCAAGCTCAGAGATGAGAGGGGAGCTCA 5522
Qy 417 gSerArgTyrGluSerSerAsnLysAlaLeu----- 427
Db 5523 GTCCAGTTCAAGTCCACTATCGCGGCTGAGGCGCAAGATGCCAGCTGAGAGCA 5582
Qy 428 ----LeuGluMetAlaGluLysThrVal-----ArgAsp 439

Db 5583 GCTTGGAGGAGGCGCAGAGAGCAAGCGCGCCCGCTGAGAGGAGCA 5642
Qy 439 sGlnLeuGluGlyLeuGlnValLysLysLysLysLysLysLysLysLysLysLysLys 459
Db 5643 GAAGCTGAAGAGATGCTGCTGAGTGAAGACGAGCGCAAGATGGCTGA-GCAGTACA 5701
Qy 459 n-----ThrGluArgAsnAspLeuAsnLysArgValGlnAspLeuSerAlaGly 476
Db 5702 AGAGCAGGACGAGAAAGAAAGCAAGCCCAAGGTCAAGCAGCTCAAGAGGAGCTGGAGGAG 5761
Qy 476 yGlnGlySerLeuThrAspSerGlyProGluArgArgProGlu----- 490
Db 5762 CCGAGGAGAGTCT-----GCAGCGCATCAAGCCCAAGCGAGGAGCTGAGCGGAGC 5815
Qy 491 -----GlyProGlyAlaGlnAlaProSerSerProArg 501
Db 5816 TGGACGAGGCGCAGGAGAGCAAGCGCGCATGGCGCGCA 5855

RESULT 26
US-08-195-487-3
Sequence 3, Application US/08195487
Patent No. 5783403
GENERAL INFORMATION:
APPLICANT: TOKATLY, GARY
APPLICANT: LIDGARD, GRAHAM P
TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE
TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA HURWITZ & THIBEAULT
STREET: 53 STATE STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/195,487
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/901,701
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: PITCHER, ESQ. EDMUND R
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: MTP-013
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/248-7000
TELEFAX: 617/248-7100
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 6306 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..6306
PUBLICATION INFORMATION:
AUTHORS: COMPTON, DUANE A
AUTHORS: SZILAK, ILLYA
AUTHORS: CLEVELAND, DON W
TITLE: PRIMARY STRUCTURE OF OF NUVA, AN INTRANUCLEAR
TITLE: PROTEIN THAT DEFINES A NOVEL PATHWAY FOR
TITLE: SEGREGATION OF PROTEINS AT MITOSIS
JOURNAL: JOURNAL OF CELL BIOLOGY
VOLUME: 116

PAGES: 1395-1408
DATE: MAR-1992
US-08-195-487-3

Alignment Scores:

Pred. No.: 3,69e-12 Length: 6306
Score: 254.00 Matches: 132
Percent Similarity: 38.42% Conservative: 87
Best Local Similarity: 23.16% Mismatches: 242
Query Match: 9.40% Indels: 109
DB: 1 Gaps: 20

US-10-023-523-8 (1-530) x US-08-195-487-3 (1-6306)

```
QY 3 SerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAla 22
DB 3714 TCGCAGGCTCTCGAGAAGAGGGGGAGAG-----3743
QY 23 ProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAla 42
DB 3744 CAAGAGGTGAAGCGGCTGTGTATGGC-----CGAGTCAGAGAGAGCCAGAGCTGGA 3797
QY 43 -----GlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSer 56
DB 3798 GGAGAGCTGCCCTGCTGCAGCAGACAGCCAGCAGAGTGCC-----AGAGCTGCA 3851
QY 57 GluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrcysValaspAsnGln 76
DB 3852 GAACGAGCTGTCTCTCGCGAGAGGCTGCAGAGCCTC-----CGGAG 3896
QY 77 GlyGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLys 96
DB 3897 CGAGCTCAGAAACAGCGGTGCTTCAGAGAACTCGCGCAGAGCTGACCTC-----3950
QY 97 SerArgThrTyrcysValaspAsnGlnGlyGluProGluProThrProValValTyrcysGlu 116
DB 3951 -----ACAGCTGAGCTGCGGAGAGAGCTGGCCAAAGATTGAAGCGTGCAGAGAGAA 4004
QY 117 LysGluProSerLysGlyAspPro-AsnThr-----126
DB 4005 GTTCTTCAGAAAGAGAGAGCGCTCTCCACCTCAGCTGAGCAGACACACAGCAGC 4064
QY 127 -----GluGluLeuArgGlnSerAspG 134
DB 4065 CCGTGTGAGTGTGCTGCGAGTAAGACCTCTGCGCAGCAGCTGCGCAGCGCAGCAGC 4124
QY 134 uValGlyAspArgAspHisArgArg-----ProGlnGluLysLysLysLysLysLys 152
DB 4125 CGCTGCCGAGAAACGCCACCTGAGAGAGCTGAGCAGCAGCAGCAGCAGCAGCAGC 4184
QY 152 uGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLys 172
DB 4185 CGCGGCGAGAGTGTGCGGGCCAGCGGAGCTGGGGAGCTGATTCCTCTGCGGCGAGAA 4244
QY 172 sLeuAlaAlaLeuCysLysLysTyrcysAlaGluLeuLeuGluGluHisArgAsnSerGlnLys 192
DB 4245 GGTGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 4304
QY 192 sGlnMetLysLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArg 212
DB 4305 CGAGCTGAGCATGTGAGAGAGCGCATGGCTGCTGCGCAGAGAGAAC-----CG 4355
QY 212 gGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLe 232
DB 4356 GGGGCTGGGTGAGCGGGCCAACTTGGCGGAGCTTCTGGAAGTG-----GAGTT 4406
QY 232 u-----GlnArgHisAsnArgSerLeuLysGluGlyValcIlnArgAlaArgGluGlu 251
DB 4407 GGACAGCGCGCGGAAGAATGATGTCAGAGATTGGCAGCGCTGCTGCTGCTGCTGCTG 4463
QY 251 uGlyLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeu 271
DB 4464 GACCCGCTGTGCTGAGGTGCGAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 4523
```

```
QY 271 nMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGlnAsnMetGluLeuAlaG 291
DB 4524 GACTGCCAAGTATGAGGTGCGCAAGGTCAAGGTCTCGGAGGAGCGCAGCGTTCACGGA 4583
QY 291 uArgLeuLysLysLeuIleGluGlnTyrcysLeuArgGluGluHisIleAspLysValPh 311
DB 4584 AGAGAGGAGAGAACTCACTGCCAGCGTGA-----4614
QY 311 elyHisLysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnAlaGlnGluMe 331
DB 4615 -----GAACCTGAGTAAGAACTGCTGACTCT-----4641
QY 331 tLeuLysGluAlaGluArgHisGlnArgGluLysAspPheLeuLysGluAlaVa 351
DB 4642 -----CACCAAGCAGCAAGGTGCAGCAGCAGAGGTGAAGCTGTCCAGGCTCAGGAGG 4697
QY 351 lGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlnThrHisLeuLysGlnGlnLe 371
DB 4698 CGAGAGCAGCAGAGGAGCGCGCTCCAGGCCAGCTGAATGAATGCAAGCCAGTT 4757
QY 371 uAlaLeuTyrcysGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluVa 391
DB 4758 GAGCCAGAGCAGCAGCAGCAGCTGAGCATAAGCTGCAGATGAGAAAGCAAAACACA 4817
QY 391 lPheThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLysLysLys 411
DB 4818 TTATGATGCCAAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 4875
QY 411 sGluThrThrMetTyrcysArgSerArgTrpGluSerSerAsnLysAlaLeuLeuGluMetAl 431
DB 4876 -----CAGCTGCAGAAAGGAAACAAAGAGCTGCAGAGCTGAAGC 4913
QY 431 aGluLysThrValArgAspLysGluLeuGluLysLysLysLysLysLysLysLysLys 451
DB 4914 TGAACGG-----CTGGCCATGAGCTACAGCAGCAGCTGGCTGGAAGCAAGAGGAGC 4964
QY 451 uGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgVal-----469
DB 4965 TGAACAGACCTGCGCGCCACCTTACTGCGCAGCTGCGCAGCTGCGGAGGACAGGTTGCCCA 5024
QY 470 -----GlnAspLeuSerAlaGlyGlyGlnGlySerLeu-----ThrAspSerGlyProG 486
DB 5025 TGCAGACAGCAGCTTCGAGACCTCGGCAATTCAGGTGCACTGATGCTTTAAAGAG 5084
QY 486 uArgArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGluAlaPr 506
DB 5085 CCGTGAGCGCCAGGCTAAG-----CCCCAGCTGAGTATTGACAGCCTGGATCTGAG 5141
QY 506 oCysTyrcysProGlyAla-----ProSerThrGluAlaSerG 518
DB 5142 CTGCGAGGAGGAGCGCCACTCAGTATCAGCAGCAAGCTGCTGCTGCTGCTGCTGCTGCTG 5201
QY 518 yGlnThrGlyProGlnGluProThrSer 527
DB 5202 CACCAAGCTGCTGAGAGAACCGCCTCA 5229
```

RESULT 27

PCT-US93-06160-3
Sequence 3, Application PC/TUS9306160
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE
INTERIOR NUCLEAR MATRIX
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESS: TESTA HURWITZ & THIBEAULT
STREET: 53 STATE STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/06160
FILING DATE: 19930621
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: FITCHER ESQ, EDMUND R
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: MTP-013
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/248-7000
TELEFAX: 617/248-7100
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 6306 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..6306
PUBLICATION INFORMATION:
AUTHORS: COMPTON, DUANE A
AUTHORS: SZILAK, ILLYA
AUTHORS: CLEVELAND, DON W
TITLE: PRIMARY STRUCTURE OF OF NUMA, AN INTRIN
TITLE: PROTEIN THAT DEFINES A NOVEL PATHWAY FO
TITLE: SEGREGATION OF PROTEINS AT MITOSIS
JOURNAL: JOURNAL OF CELL BIOLOGY
VOLUME: 116
PAGES: 1395-1408
DATE: MAR-1992
PCT-US93-06160-3

Alignment Scores:	3.69E-12	Length:	6308
Pred. No.:		Matches:	132
Score:	254.00	Conservative:	87
Percent Similarity:	38.4%	Mismatches:	242
Best Local Similarity:	23.1%	Indels:	109
Query Match:	9.40%	Gaps:	20
DB:	5		

US-10-023-523-8 (1-530) X PCT-US93-06160-3 (1-6306)

Qy	3	SerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAla	22
Db	3714	TCGCCAGGCTCTGTGAGAAGGAGGGGAGAG	3743
Qy	23	ProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAla	42
Db	3744	CAAGGAGTTGAAGCGCTGTGTATGGC-----CGAGTCAGAGAAGAGCAGGAAGCTGGA	3797
Qy	43	-----GlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSer	56
Db	3798	GGAGAGCTGGCCCTGCTGCAGGCGAGACAGCACCACAGTCCC-----AGAGCTGCA	3851
Qy	57	GluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGln	76
Db	3852	GAACGACGCTCTGCTCTCGGGAGGAGGTGCAGAGCCTC-----CGGGAG	3896
Qy	77	GlyGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLys	96
Db	3897	GGAGGCTGGAACACGCGCGGTGGCTTCAGAGAACCCTCGCGAGAGCTGACCTC	3950
Qy	97	SerArgThrTyrValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGlu	116
Db	3951	-----ACAGGCTGAGCGTGGGAGGAGCTGGGGCAAGATTGAGCGCTGGCAGGAA	4004
Qy	117	LysGluProSerLysGlyAspPro-AasnThr-----	126

Db 4965 TGAACAGACCTGCCGCCACCTTACTGCCAGGTGCGCAGCTGGAGSCACAGGTTGCCCA 5024

Qy 470 -----GlnAspLeuSerIaGlyGlyGlnGlySerLeu---ThrAspSerGlyProG 486

Db 5025 TGCAGACAGCAGCAGCTTCGAGACCTGGGCAAAATTCAGGTGGCAACTGATCTTTAAAGAG 5084

Qy 486 uArgArgProGlyGlyProGlyAlaGlnAlaProSerSerProArgValThrGluAlaPr 506

Db 5085 CCGTGAGCCCGAGCTAAG-----CCCGAGCTGGACTTGAGTATTGACAGCCTGGATCTGAG 5141

Qy 506 oCysTyProGlyAla-----ProSerThrGluAlaSerGI 518

Db 5142 CTGGAGAGGGGAGCCCACTCACTAGTATCACCAGCAAGCTGCTCGTACCAGCCAGCAGCG 5201

Qy 518 yGlnThrGlyProGlnGluProThrSer 527

Db 5202 CACCAGCGTCCCTGGAGAACCGCTCA 5229

RESULT 28

US-08-466-390-3

Sequence 3, Application US/08466390

Patent No. 5686562

GENERAL INFORMATION:

APPLICANT: TOURATLY, GARY

INVENTOR: LIDGARD, GRAHAM P

TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE

TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: TESTA, HORWITZ & THIBEAULT

STREET: 125 HIGH STREET

CITY: BOSTON

STATE: MA

COUNTRY: USA

ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/466,390

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: PITCHER ESQ, EDMUND R

REGISTRATION NUMBER: 27,829

REFERENCE/DOCKET NUMBER: MTP-013

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 248-7000

TELEFAX: (617) 248-7100

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 6306 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 1..6306

PUBLICATION INFORMATION:

AUTHORS: COMPTON, DUANE A

AUTHORS: SZILAK, ILYA

AUTHORS: CLEVELAND, DON W

TITLE: PRIMARY STRUCTURE OF NUMA, AN INTRANUCLEAR

TITLE: PROTEIN THAT DEFINES A NOVEL PATHWAY FOR

TITLE: SEGREGATION OF PROTEINS AT MITOSIS

JOURNAL: J. Cell Biol.

VOLUME: 116

PAGES: 1395-1408

DATE: 1992

US-08-466-390-3

Alignment Scores:

Pred. No.: 4,52e-12 Length: 6306

Score: 253.00 Matches: 132

Percent Similarity: 38.42% Conservative: 87

Best Local Similarity: 23.16% Mismatches: 242

Query Match: 9.36% Indels: 109

DB: 1 Gaps: 20

US-10-023-523-8 (1-530) x US-08-466-390-3 (1-6306)

Qy 3 SerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAala 22

Db 3714 TCGCAGCTCTGGAGAGAGAGGGGGAGAG----- 3743

Qy 23 ProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGlyGlyAla 42

Db 3744 CAAGGAGTTGAAGCGGTGGTATGGC-----CGAGTCAGAGAGAGCCCAAGAGCTGGA 3797

Qy 43 -----GlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValser 56

Db 3798 GGAGAGCTGCGCTGCTGCAGGAGAGACAGCCAGCAACAGTGCC-----AGAGCTGCA 3851

Qy 57 GluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnGln 76

Db 3852 GAACGAGCTCTGCTCTCGGGAGGAGGTGCAGAGCCTC-----CGGAG 3896

Qy 77 GlyGlyProGlyGluAspGlyAlaGlnGlyProAlaGluProGluAspAlaGluLys 96

Db 3897 GGAGCTGAGAAACAGCGGTTGGCTTCAGAGAACTCGGCGAGGAGTGACTC----- 3950

Qy 97 SerArgThrTyrValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGlu 116

Db 3951 -----ACAGGCTGAGCGTGGCGAGGAGCTGGGCCAAGATTGAAGCGCTGGCAGAGAA 4004

Qy 117 LysGluProSerLysGlyAspPro-AsnThr----- 126

Db 4005 GTTCTTCAGAAAGAGAGAGCGCCCTCTCCACCTCGAGCTCGAGCACACCAGCACAGGC 4064

Qy 127 -----GluGluIleArgGlnSerAspGl 134

Db 4065 CCTGCTGAGTGAGCTGCTGCCAGCTTAAGCACCTCTGCCAGCAGCTGCGCGCAGCAGGC 4124

Qy 134 uValGlyAspArgAspHisArgArg-----ProGlnGluLysLysLysAlaLysGlyLe 152

Db 4125 CGCTGCCAGAGAACGCCACCTGAGGAGCTGGAGCAGAGCAAGCAGCCCTGGGGGACT 4184

Qy 152 uGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLys 172

Db 4185 GCGGSCAGAGCTGCTCGCGGCGCCAGCGGAGCTTGGGGAGCTGATTCTCTCGCGCAGAA 4244

Qy 172 sleuAlaAlaLeuCysLysLysTyrAlaGluLeuLeuGluHisArgAsnSerGlnLys 192

Db 4245 GGTGGCAGAGCAGGAGAGCAACAGCTTCAGAGCTCGGGCAGAGAGAGCCAGCTATGAGA 4304

Qy 192 sGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuAr 212

Db 4305 GCAGCTGAGCATGTGAAGAGCGCGCATGGCTGCTGGCAGAGAGAAC-----CG 4355

Qy 212 gGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLe 232

Db 4356 GGGGCTGGGTGAGCGCGGCCAACCTTGGCGCGCAGTTCTTGGAAAGTG-----GAGTT 4406

Qy 232 u--GlnArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGluGlu 251

Db 4407 GGACCGAGCCCGGGAAAGATGTCCAGAGATTGGCAGCCCGTACGTCT---GATGCTGA 4463

Qy 251 uGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGl 271

Db 4464 GACCCGCTCTGGCTGAGGTGCAGCGAGAGACAGACAGCAGCTCCCGGGAGGTGGAGTGA 4523

Qy 271 nMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGlnGluAsnMetGluLeuAlaGl 291

4005 GTTCTTCAGAAAGAGCAGCGCCCTCTCCACCTGAGCTCGAGCACACACAGCAGGC 4064
127 -----GluGluLeuArgGlnSerAspG1 134
4065 CTTGTGTAGTGTGCTGCTGCGAGCTAAGCACCTCTGCGCAGCAGCTGCGCGCAGCAGGC 4124
134 uValGluAspArgAspHisArgArg-----ProGlnGluLysLysLysLysLysLysLys 152
4125 CGTTCGCGAAGAACGCCACCTGAGGAGCTGGAGCAGAGCAGCAGCGCTGGGGGACT 4184
152 uGluLysGluLeuThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLys 172
4185 CGCGCAGAGCTGTGCGGGGCCAGCGGAGCTTGGGGAGCTGATCTCTGCGCGCAGAA 4244
172 sLeuAlaAlaLeuCysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 192
4245 GGTGGCAGAGCAGAGCGAGCGAAGCAGCTCAGCAGCTCGCGGCGAGAGCGCCTATGCGA 4304
192 sGlnMetLysLeuLeuGlnLysLysLysLysLysLysLysLysLysLysLysLysLysLys 212
4305 GCAGCTGAGCATGCTGAAGAGCGCGCATGCGCTCTGCGCAGAGAGAAC-----CG 4355
212 gGluGluHisSerLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 232
4356 GGGCTGGTGGAGCGCGCCACCTTGGCGCGCAGTTCTGGAAGTG-----GAGTT 4406
232 u---GlnArgHisAsnArgSerLeuLysGluGluGluValGlnArgAlaArgGluGlu 251
4407 GGACCGCGCGGAAAGATGATGCAAGAGTTGGCAGCGCTGAGCTGCT---GATCTGA 4463
251 uGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspGlnLeuG1 271
4464 GACCGCTGCTGAGTGTGAGCGGAGCAGCAGCAGCAGCTGCCCGGAGCTGAGGTGAT 4523
271 nMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGlnGlnGlnGlnGlnGlnGln 291
4524 GACTGCCAAGTATGAGGGTCCCAAGGTCAAGGTCTCTGGAGGAGCAGCAGCGGTTCAGGA 4583
291 uArgLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 311
4584 AGAGAGCAGAAACTCACTCCCGAGTGGAA----- 4614
311 eLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluWe 331
4615 -----GAACTGAGTAAGAACTGGCTGACTCT----- 4641
331 tLeuLysGluAlaGluGluArgHisGlnArgGluLysAspPheLeuLysGluAlaVa 351
4642 ----GACCAAGCAGCAGAGTGCAGCAGCAGCAGCAGTGAAGGTGCTCCAGGCTCAGGAGG 4697
351 lGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlnThrHisLeuLysGlnGlnLe 371
4698 CGAGAGCCAGCAGGAGCGCCCGCGCTTCCAGGCCAGCTGAATGAATCAGAGCCAGTT 4757
371 uAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluVa 391
4758 GACCCAGAGAGCAGCAGCTGAGCTGAGCTATGAGTGCAGATGGAGAGGCAACCAACACA 4817
391 lPheThrThrPheLysGlnGlnMetGluLysMetThrLysLysLysLysLysLysLysLys 411
4818 TTATGATGCCAAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 4875
411 sGluThrThrMetTyrArgSerArgTrpGluSerSerAsnLysAlaLeuLysGluVal 431
4876 -----CAGCTGCAGAGAGGAAACAAAGAGCTCGAGCTCGAGCTGAAGC 4913
431 aGluGluLysThrValArgAspLysGluGluGluGluGlnValLysLysLysLysLysLys 451
4914 TGAACGG-----CTGGGCCATGAGCTACAGCAGGCTGGGCTGAGAGCAACAGAGGC 4964
451 uGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgVal----- 469
4965 TGAACAGACCTCGCGCCACCTTACTGCCAGGTGCGCAGCTGGAGGACAGAGTTGCCCA 5024

470 -----GlnAspLeuSerAlaGlyGlyGlnGlySerLeu---ThrAspSerGlyProG1 486
5025 TGCAGACAGCAGCAGCTTCGAGACCTGGGCAAAATTCAGAGTGGCACTGATGCTTTAAAGAG 5084
486 uArgArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGluAlaPr 506
5085 CCGTGAGCGCCCGAGCTTAAG---CCCCAGCTGGACTTGATTTGACAGCCTGGATCTGAG 5141
506 oCysTyrProGlyAla-----ProSerThrGluAlaSerG1 518
5142 CTGCGAGGAGGGGACCCCACTCAGTATCACCAGCAGCTGCTCGTACCCAGCAGCAGCG 5201
518 yGlnThrGlyProGlnGluProThrSer 527
5202 CACCAGCGTCCCTGGAGAACCCAGCCTCA 5229
RESULT 30
US-08-467-781-3
Sequence 3, Application US/08467781
Patent No. 5780596
GENERAL INFORMATION:
APPLICANT: TOUTKATLY, GARY
APPLICANT: LIDGARD, GRAHAM P
TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE
TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: TESTA, HURWITZ & THIBEAULT
STREET: 125 HIGH STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,781
FILING DATE: 06-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: PITCHER ESO, EDMUND R
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: MTP-013
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 6306 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
FEATURE:
NAME/KEY: CDS
LOCATION: 1..6306
PUBLICATION INFORMATION:
AUTHORS: COMPTON, DUANE A
AUTHORS: SZILAK, ILYA
AUTHORS: CLEVELAND, DON W
TITLE: PRIMARY STRUCTURE OF NUMA, AN INTRANUCLEAR
TITLE: PROTEIN THAT DEFINES A NOVEL PATHWAY FOR
TITLE: SEGREGATION OF PROTEINS AT MITOSIS
JOURNAL: J. Cell Biol.
VOLUME: 116
PAGES: 1395-1408
DATE: 1992
US-08-467-781-3

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/483,924
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: PITCHER ESO, EDMUND R
 REGISTRATION NUMBER: 27,829
 REFERENCE/DOCKET NUMBER: MTP-013
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 248-7000
 TELEFAX: (617) 248-7100
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 6306 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..6306
 PUBLICATION INFORMATION:
 AUTHORS: COMPTON, DUANE A
 AUTHORS: SZILAK, ILVA
 AUTHORS: CLEVELAND, DON W
 TITLE: PRIMARY STRUCTURE OF NIMA, AN INTRANUCLEAR
 TITLE: PROTEIN THAT DEFINES A NOVEL PATHWAY FOR
 TITLE: SEGREGATION OF PROTEINS AT MITOSIS
 JOURNAL: J. Cell Biol.
 VOLUME: 116
 PAGES: 1395-1408
 DATE: 1992
 US-08-483-924-3

Alignment Scores:
 Pred. No.: 4,52e-12 Length: 6306
 Score: 253.00 Matches: 132
 Percent Similarity: 38.42% Conservative: 87
 Best Local Similarity: 23.16% Mismatches: 242
 Query Match: 9.36% Indels: 109
 DB: 2 Gaps: 20

US-10-023-523-8 (1-530) x US-08-483-924-3 (1-6306)

QY 3 SerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAla 22
 Db 3714 TCGCCAGGTCTCGAGAGGAGGGGGAGAG-----3743
 QY 23 ProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAla 42
 Db 3744 CAAGGAGTTGACGGCGTGTGTATGCC-----CGAGTCAGAGAAGAGCCAGAGCTGGA 3797
 QY 43 -----GlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSer 56
 Db 3798 GGAGAGTCGCCCTGTGAGGAGGAGAGCAGCAGCAGCAGTGC-----AGAGCTGCA 3851
 QY 57 GluGluLeuSerArgGlnLeuGluAspIleLeuSerThrThrCysValAspAsnGln 76
 Db 3852 GAACGCGAGTCTCTCTCGGGAGGAGGTGCAGAGCCTC-----CGGGAG 3896
 QY 77 GlyGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLys 96
 Db 3897 GGAGGCTGAACACAGCGGTGCTTCAGAGAACCTGCGGCGAGGAGCTGACCTC-----3950
 QY 97 SerArgThrTyrValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGlu 116
 Db 3951 -----ACAGGCTGAGCGTGGGAGGAGCTGGGCCCAAGAAATTGAAGCGGTGCGAGAGAA 4004
 QY 117 LysGluProSerLysGlyAspProAsnThr-----126
 Db 4005 GTTCTTCCAGAAAGACGAGCGCCCTCTCCACCTGTGAGCTGAGAGCACACGAGCACAGGC 4064

QY 127 -----GluGluIleArgGlnSerAspG1 134
 Db 4065 CTTGGTGTAGTGTGCTGCCAGCTTAAGCACCTCTGCCAGCAGCTGTCAGGCGGAGCAGGC 4124
 QY 134 uValGlyArgAspHisArgArg-----ProGlnGluLysLysLysAlaLysGlyLe 152
 Db 4125 CCTGCCGAGAAACCCACCGTGTGAGGAGCTGAGCAGAGCAGGAGCGCTGGGGGACT 4184
 QY 152 uGlyLysGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLys 172
 Db 4185 GCGGCGCAGAGTGTCTCGCGGCCAGCGGAGCTTGGGAGCTGATTCCTCTGCGGCAGAA 4244
 QY 172 sleuAlaLeuLysLysLysTyrAlaGluLeuLeuGluHisArgAsnSerGlnLys 192
 Db 4245 GGTGCGCAGCAGGAGCGAACAGCTCAGCAGCTGCGGCGCAGAGGCGCAGCTATGCAGA 4304
 QY 192 sGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuAr 212
 Db 4305 GCAGCTGAGCATGCTGAAGAAGCGCGCATGCTGTCGAGAGGAGAAC-----CG 4355
 QY 212 sGlyGluHisSerLysAlaValLeuAlaArgSerLysLysLeuGluSerLeuCysArgGluLe 232
 Db 4356 GGGGCTGGGTGAGCGCGGCAACCTTGGCGCGCAGTTTCTGGAAGTG-----GAGTT 4406
 QY 232 u---GlnArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGluGlu 251
 Db 4407 GGACCGAGCGCGGAAAGTATGTCAGAGATTGGCAGCCCTAGCTGCT---GATGCTGA 4463
 QY 251 uGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuG1 271
 Db 4464 GACCCGTCTGCTGAGGTGCAGCGAGAGACAGACAGCAGCTGCCCGGAGCTGGAGGTGAT 4523
 QY 271 nMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGlnAsnMetGluLeuAlaG1 291
 Db 4524 GACTCCCAAGTATGAGGTGCCAAGTCAAGGTCTCTGGAGGAGGAGCGCGGTTCACGA 4583
 QY 291 uArgLeuLysLysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAspLysValPh 311
 Db 4584 AGAGAGGCGAGAACTCACTGCCAGGTGGA-----4614
 QY 311 eLysHisLysAspLeuGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMe 331
 Db 4615 -----GAAGTGAAGTGAAGAACTGGCTGACTCT-----4641
 QY 331 tLeuLysGluAlaGluArgHisGlnArgGluLysAspPheLeuLysGluAlaLava 351
 Db 4642 ---GACCAGCCAGCAGGTGTCAGCAGCAGAGCTGAAGGCTCTCCAGGCTCAGGAGG 4697
 QY 351 lGluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLysLysGlnGlnLe 371
 Db 4698 CGAGAGCAGCAGAGGAGCGCCAGCGCTTCCAGGCGCAGCTGAATGAATGAATGAATGAAT 4757
 QY 371 uAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLysSerSerGluVa 391
 Db 4758 GAGCCAGAGGAGCAGCAGCTGAGCAGCTATAGCTCAGATGAGAAAGCCAAACACACA 4817
 QY 391 lPheThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLysLysLys 411
 Db 4818 TTATGATGCCAAGAGCAGCAGCAACCAAGAGCTGTCAGGAGCAGCTGGCGGAGCTGGAG-- 4875
 QY 411 sGluThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeuMetal 431
 Db 4876 -----CAGTCGAGAGAGGAAACAAAGAGCTGCGAGCTGAAGC 4913
 QY 431 aGluGluLysThrValArgAspLysGluLeuGluGlyLeuGlnValLysIleGlnArgLe 451
 Db 4914 TGAACGG-----CTGGGCGCATGAGCTACAGAGCTGGGCTGAAGCAAGCAGGAGGC 4964
 QY 451 uGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgVal----- 469
 Db 4965 TGAACAGACAGCTGCGGCCACCTTACTGCCCCAGGTGCGCAGCTGAGGAGCAGAGTTGCCCA 5024

QY 470 -----GlnAspLeuSerAlaGlyGlyGlnGlySerLeu---ThrAspSerGlyProG1 486
 Db 5025 TCAGACAGCAGCTTCAGACCTCGGCATATTCAGGTGGCACTGATGCTTTAAAGAG 5084
 QY 486 uArgArgProGluGlyProGlyAlaGlnAlaProSerSerProArgValThrGluAlaPr 506
 Db 5085 CCGTGAGCCCCAGGCTAAG---CCCAGCTGGACTTGAGTATTGACACGCTGGATCTGAG 5141
 QY 506 oCysTyrProGlyAla-----ProSerThrGluAlaSerG1 518
 Db 5142 CTGCGAGGAGGGAGCCCACTCACTAGTATCACACAGAGTGGCTGCTCATCCAGCCAGAGCG 5201
 QY 518 yGlnThrGlyProGlnGluProThrSer 527
 Db 5202 CACCGAGCTCCCTGGAGAACCACTCA 5229

RESULT 32

US-09-976-594-640
 ; Sequence 640, Application US/09976594
 ; Patent No. 6673549
 ; GENERAL INFORMATION:
 ; APPLICANT: Furness, Michael
 ; APPLICANT: Buchinder, Jenny
 ; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
 ; FILE REFERENCE: PA-0041 US
 ; CURRENT APPLICATION NUMBER: US/09/976,594
 ; CURRENT FILING DATE: 2001-10-12
 ; PRIOR APPLICATION NUMBER: 60/240,409
 ; PRIOR FILING DATE: 2000-10-12
 ; NUMBER OF SEQ ID NOS: 1143
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 640
 ; LENGTH: 5185
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; OTHER INFORMATION: Incyte ID No. 6673549 346209.3
 US-09-976-594-640

Alignment Scores:
 Pred. No.: 5 62e-12 Length: 5185
 Score: 250.50 Matches: 140
 Percent Similarity: 38.27% Conservative: 95
 Best Local Similarity: 22.80% Mismatches: 232
 Indels: 150
 Gaps: 24
 DB:

US-10-023-523-8 (1-530) x US-09-976-594-640 (1-5185)

QY 5 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaAlaProAla 24
 Db 611 GGQCATCCGCATCTGTGCGCAGGGCTTCCCAACCGCATCTCTTCCAGGAGTTCGGCA 670
 QY 25 Val-----GluAla--- 27
 Db 671 GCGATACGATCTGACACCCCAATGCCATCCCAAGGGCTTCATGGATGGAGGAGCGC 730
 QY 28 -----GluGlyProGlySerSerGlnAlaProArgLys---ProGluGlyAlaGln 43
 Db 731 CTGTGAAAGATGATCCAGCGCTGGAACTGGACCCCAACCTCTACCGGTGGGACAGAG 790
 QY 44 AlaArgThrAla-GlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLe 63
 Db 791 CAAGATCTTCTCCGGCTGGGTCTGGCCAGCTGGAAAGAGGAGGAGACCTGAGGT 850
 QY 63 uGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyProGlyGluAspG1 83
 Db 851 CACGACATCATCTCTCTTC-----CAGCGAGCTGC-CCGGGGATACCTGG 897
 QY 83 YAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrValAlaAr 103
 Db 898 CTCGAGGGCTTCCAGAGCGCCAGCAGCAGCAGCGCCCTGAGGGTGATGACGGGA 957

QY 103 g---AsnGlyGluProGluPro-----ThrProValValTyrGlyGluY 117
 Db 958 ACTGCGCGGCTACTCAAGCTGAGACACTGGCAGTGGTGGCGGCTGTTTACCAAGTGA 1017
 QY 117 sGluProSerLysGlyAspProAsnThrGluGluLeuArgGlnSerAspGluValGlyAs 137
 Db 1018 AGCCACTGCTGCAAGGTGACG--CGGCAGGATGAGTGGCTGCAGGCA----- 1061
 QY 137 pArgAspHisArgProGlnGluLysLysAlaLysGlyLeuGlyLysGluLeuTh 157
 Db 1062 -----CGGCCCGAGGAGCTGCAGAAAGTGCAGAGGCTACAGCAGCAG--- 1103
 QY 157 rLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCy 177
 Db 1104 -----AGCGCCCGCGAAGTTGGGAGGCTCCA 1129
 QY 177 sLysLysTyrAlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeu 197
 Db 1130 GGGCGGAGTGGCAGCAGCTGGAAGAGGAGCGCGCGCGCTGGCAGAGCAATG--- 1181
 QY 197 uGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLy 217
 Db 1182 ----CGAGCAGAGCGCAGCACTGTGTGCAGAGCGCGCAGAGACGCGGG---AG 1228
 QY 217 sAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnAr 237
 Db 1229 GCTGCGCAGCGCGCAAGCAGGAGCTGGAGCTGGTGTCTCAGAGCTGGAGCT---CG 1282
 QY 237 gSerLeuLysGluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluVa 257
 Db 1283 CGTGGCGGAGGAGGAGGAGTGCACCCCTCAATGCACACCGCAGAGAGAGG--- 1334
 QY 257 lThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGlnGlnHisAsnG1 277
 Db 1335 -----CTACAGCAGCAGCATACAGGAGCTAGAGCGCCACCTCAGGCTGAGGAGGG 1384
 QY 277 uArgAsnSerLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuI1 297
 Db 1385 TGCGCGGCGCAAGCTGCACTGGAGAGGTGACACAGAGGCAAAATGAGAAATTGA 1444
 QY 297 eGluGlnTyrGluLeuArgGluGluHisIleAspLysValPheLys---HisLysAspLe 316
 Db 1445 AGAGGACCTGTCTCTCTGGAAGACCAAGATTCACAGCTGACGAGGAGCGGAGCTGCT 1504
 QY 316 uGlnGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGlu----- 330
 Db 1505 GGAAGATCGTCTGCGCGAGTTCTCATCCAGCGAGCTGAGGAGGAGGAGAGGTCAGAG 1564
 QY 331 -----MetLe 332
 Db 1565 CCTCAATAAGCTACGGCTCAAATATGAGGCCACAATCGCAGACATGGAGGACCGCTACG 1624
 QY 332 uLysGluAlaGluGluArgHisGlnArgGlyLysAspPheLeuLysGluAlaVal-- 351
 Db 1625 GAAAGAGGAGAGGGTCCCGCAGGAGCTGGAGAG-----CTGAAAGCGGAGCTGGA 1675
 QY 352 ----GluSerGlnArgMetCysGluLeuMetLysGlnGlnGlnThrHisLeuLysGlnG1 370
 Db 1676 TCGGGAGAGCTCAGAGCTGCGAGGAGCAGATGTTGGAGCAGCAACAGCGCGCAGAGGAGCT 1735
 QY 370 nLeuAlaLeuTyrThrGluLysPheGluPheGlnAsnThrLeuSerLysSerSerG1 390
 Db 1736 GCGGCGCCAGCTGGCGCGAAGAGGAGGAGCTGCAGGCTGCCCTGGCGCAGGAGAGA 1795
 QY 390 uValPheThrThrPheLysGlnGlnMetGluLysMetThrLysLysIleLysLysLeu-- 409
 Db 1796 C-----GAGGTGGGCGCGCGCGCCAGCTGCTGGAATCCCTCGC 1834
 QY 410 -GluLysGluThrThrMetTyrArgSerArgTrpGluSerSerAsnLysAlaLeuGluG1 429
 Db 1835 GGAGGCTCAAGCAGCGCTGGCGCGAGGCCAGGAGGACCTGGAGTCTGAGCGTGTGGCCAG 1894

```

Qy 429 uMetAlaGluGlyLysThrVala:qAsp-----LysGluLeuGluGlyLeuGlnVally 447
Db 1895 GACCAAGCGGAGAGAGCAGCGCCGGACCTGGGGCAGAGCTGGAGCGCTGGGGCGGA 1954
Qy 447 sile-----GL 449
Db 1955 GCTGAGGACACGCTGGACTCCACCAACGCACAGCAGGAGCTCCGGTCCAGAGGGAACA 2014
Qy 449 nArgLeuGluLysLeuGlnGlnThrGluArgAsnAspLeuAsnLysArgVa 469
Db 2015 GCAGTACGAGGAGTGAAGAGACTCTGGAGGAGAGACTCGCATCCAGAGCGCGCAGT 2074
Qy 469 lInAspLeuSerAla-----GlyGlyGlnGlySerLeuThrAspse 483
Db 2075 GCAGGAGCTGAG-GCAGCGCCAGCGCCAGCGCCCTGGGGGAGCTGGCGGAGCAGCTGGAGC 2133
Qy 483 xGlyProGluArg-----ProGluGlyProGlyAlaGlnAlaPr 497
Db 2134 AGGCCCGAGGGGCAAGAGTGCATGGAGAGAGACCCCGGCTGGGCCCTGGAGCGCGAGTGT 2193
Qy 497 oSerSerProArgValThrGluAlaPro-----CysTy:ProGlyAla 511
Db 2194 CCGAGCTCGCGGCAAGACTGCGAGCGCTCGAGACTGCACGTGAGGAGCGGTGACGCGGGA 2253
Qy 512 -----ProSerThrGluAlaSerGlyGlnThrGlyPro 522
Db 2254 GCGCGCCGCTGGAGTTACAGCTGCGAGGAGGTGCGAGGCGCG 2293

RESULT 33
US-09-688-188B-20
; Sequence 20, Application US/09688188B
; Patent No. 656716
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHYTE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; FILE REFERENCE: 038602/0328 US/09/688,188B
; CURRENT APPLICATION NUMBER: 2000-10-16
; PRIOR FILING DATE: 1999-04-14
; PRIOR FILING DATE: 1999-04-14
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 20
; LENGTH: 3824
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-688-188B-20

Alignment Scores:
Pred. No.: 5,38e-12 Length: 3824
Score: 248.50 Matches: 139
Percent Similarity: 39.75% Conservative: 117
Best Local Similarity: 21.58% Mismatches: 197
Query Match: 9.20% Indels: 191
DB: 4 Gaps: 29

US-10-023-523-8 (1-530) x US-09-688-188B-20 (1-3824)
Qy 23 ProAlaValGluAlaGlu----- 28
Db 326 CCAGCAGTAGACACACAGGAGAGAGAGACAGATCATGTTGTCGCGGACGGA 385
Qy 29 -----GlyProGlySerGlnAla-----ProArgLysProGluGlyAlaGlnAla 44
Db 386 ACAGTTTAATAGTGTGGAAAGTAAATCAATCCATCCCGACATGTCATCCAGTCCGACGAGC 445
Qy 45 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 64
Db 446 CAAGCAGTAGTGTAAACAGTCTTCAGATGTCTCAGATGACAAAGT---GAGCTAGAC 502

```

```

Qy 65 AspileLeuSerThrTy:Ty:ValAspAsnAsnGlnGly-----GlyProGly 80
Db 503 ATGATGGAGGGAGACACACAGTGTATCTTACAGTCTCTGTATCCATTTAAACACGAG 562
Qy 81 GluAspGly-----AlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArg 98
Db 563 GAAGAAATATACAGAGAAGAGGAGATCCT-----AGA 595
Qy 99 ThrTy:ValAlaArgAsnGlyGluProGluProThrProValValTy:GlyGluLysGlu 118
Db 596 ACA-----AGAGCATCAGATCCACATCTCCACCCCAAGTATCTGTCACAAATCA 646
Qy 119 ProSerLysGlyAspProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArg 138
Db 647 CACTATCGTAATCGAGACACTTTGCTACTATACGACAGCATCACTGGTT----- 697
Qy 139 AspHisArgArgProGlnGlyLysLysAlaLysGlyLysGlyLysGluLeuThr--- 157
Db 698 ---ACGAGCAAAATGCAAGAACATGAGCAGACTCTGAGCTTAGAGAACAAATCTCTGGC 754
Qy 158 -----LeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAla 174
Db 755 TATAGCGAATGAGCGGACAAATCAAAAGCAACTGATGACTCTGGAAAAACAAGCTAAAG 814
Qy 175 AlaLeuCysLysLysTy:Ala-----GluLeuLeuGluGluHisArgAsn--- 189
Db 815 GCTGAGATGGATGAACATCGCCTCAGATTAGACAAAGATCTTGAACACTCAGCGTAAACAAT 874
Qy 190 ---SerGlnLysGlnMetLysLeuLeuGlnLysGlnSerGlnLeuValGlnGlu--- 207
Db 875 TTGCTGCGAAATGCGAGAACTTATCAAGAAACACCCAGGCTGCCATGGAGAAAGAGGCT 934
Qy 208 -----LysAspHisLeuArgGlyGluHisSerLys 217
Db 935 AAAGTGTGTCATGAAGAGAAAAAATTTACAGCAACATATTACAGGCCCAACAGAGAAA 994
Qy 218 AlaValLeuAlaArgSerLysLeuGluSerLysCysArgGluLeuGlnArgHisAspArg 237
Db 995 GAACGTG-----AATAGTTTTCTCGAGTCCAGAAAAAGAGAGATATAAATCTCGAAAAAGAG 1048
Qy 238 SerLeuLysGluGluGlyValGlnArg-----AlaArgGluGluGluGlu--- 252
Db 1049 CAGCTTAAGAGGAGCTAAATGAACACAGAGTACCCCAAAAAAGAAAAACAGAGATGG 1108
Qy 253 -----LysArgLysGluValThrSerHisPheGlnVal-----ThrLeuAsnAspIle 268
Db 1109 CTTTCAAGCAGAGAGAGATATACAGCAATTTCCAGCAGAGAGAGAGAGTAACTTCTT 1168
Qy 269 GlnLeuGlnMetGlu----- 273
Db 1169 CGAGCTCAAGACAAATACCTAGAGCTGGAATGCGCTTCAAGAGAGAGAGAGTACTT 1228
Qy 274 ---GlnHisAsn-----GluArgAsnSerLysLeuArgGln 284
Db 1229 GGGGTCTAATCTAGAGCAGGAGCTTGTGAGGAGAGAGTAAACAAAGACAGAGCTCAG 1288
Qy 285 GluAsnMetGluLeuAlaGluArgLeuLysLeuLeuGlnTy:GluLeuArgGlu 304
Db 1289 AAGGACTTAGAGCATGCCATGCTACTCCGACAGCATGAATCTATGCAAGAGAGAGAGTTC 1348
Qy 305 GluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeuValAspAlaLys 324
Db 1349 CGCCACCTCAACAAATTCAGAAG-----ATCGCTGTGAGTTGATC-----AGA 1393
Qy 325 LeuGlnGlnAlaGlnGlnMetLeuLysGluAlaGluGlnArgHisGlnArgGluLysAsp 344
Db 1394 TTACAGCATCAAACTGAGCTCACTAACCCAGCTGGAATATAATAAGCGAGAGAGACGAGAA 1453
Qy 345 PheLeuLeuLysGluAlaValGlnSerGlnArgMetCysGluLeuMetLysGlnGlnGlu 364
Db 1454 CTAAGACAAAGCATGCTCAGTGAAGAGTTCGACACAGCCTAAGAGTTGAGTGAAGAA 1513

```



```

QY 365 ThrHisLeuLysGlnGlnLeu-----AlaLeuTyrThrGluLysPheGluGlu 380
Db 1514 CTCCAATATAAAGACGCTTTTCAGGATACCTGCAAAATCCAAACACAGACATACAAAGCA 1573
QY 381 PheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGlu 400
Db 1574 TTAAGNAATCCTCTCTGGAG-----ACTACACCAAGAGTGGACACAAA 1618
QY 401 LysMetThrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 420
Db 1619 GCTGTTCTTGAACCGCTCAAG-----GAGGAACACAGCCCGGAATTAGCTATCTTGCTG 1672
QY 421 GluSerSerAsnLysAlaLeuLeuLeuMetAlaGluGluLysThrValArg----- 437
Db 1673 GAGCAGTATGATCAGACGATTAATGAAATGCTCTCCACACAGCCCTGGTTGGATGAA 1732
QY 437 ----- 437
Db 1733 GCACAGGAGCAGAGTGCAGGCTTTTGAAGATGCAGCTGCAGCAGGAACTGGAGCTGTG 1792
QY 438 ----- 438
Db 1793 AATGCGTATCAGAGCAAAATCAAGATGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1852
QY 443 GlyLeuGlnValLysLys-----GlnArgLeuGluLysLeu 454
Db 1853 GAGCTTGAACAGAGGCTCTCCCTCCGAGGGCAGCTCTTAGAACAAGATTGAAGAGAG 1912
QY 455 CysArgAlaLeuGlnThrGluArgAsnAsp-----LeuAsnLysArgVal 469
Db 1913 ATGTTGGCTTTTGCAGAAATGAGCGCAGACGAATACGAAGCGCTGTGGAAGCTCAAGCC 1972
QY 470 GlnAspLeuSerAlaGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 480
Db 1973 AGAGAGATTGAAGCTTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2032
QY 481 ThrAspSerGlyProGluArgArgProGluGlyProGlyAlaGlnAlaProSerSerPro 500
Db 2033 TCTAATCTCTCCCTGAG----- 2050
QY 501 ArgValThrGluAlaProCysTyrProGlyAlaPro-----SerThrGluAlaSerGly 518
Db 2051 -----GCATTTCAGCAGCAGCTACCCGGAGCTTCTGGTTGTTGTCACAAACCTACTGG 2104
QY 519 GlnThrGlyPro 522
Db 2105 GGTCCAGGACCT 2116

RESULT 34
US-09-291-417D-20
; Sequence 20, Application US/09291417D
; Patent No. 6680170
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHITE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; FILE REFERENCE: 038602/0329
; CURRENT APPLICATION NUMBER: US/09/291,417D
; CURRENT FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 60/081,784
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 3824
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-291-417D-20

```

Alignment Scores: 5.38e-12 Length: 3824
Pred. No.: 248.50 Matches: 139
Score:

```

Percent Similarity: 39.75% Conservativity: 117
Best Local Similarity: 21.58% Mismatches: 197
Query Match: 9.20% Indels: 191
DB: Gaps: 29

US-10-023-523-8 (1-530) x US-09-291-417D-20 (1-3824)

QY 23 ProAlaValGluAlaGlu----- 28
Db 326 CCAGCATGAAGACACAGGAGGAAGAACAGATCATGGTGTGGCCGACAGGA 385
QY 29 -----GlyProGlySerSerGlnAla-----ProArgLysProGluGlyAlaGlnAla 44
Db 386 ACAGTTAATAGTGTGGAGTAAATCAATCCATTCCTCCAGCATGTCCATCACTGCGCAGCAGC 445
QY 45 ArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeuGlu 64
Db 446 CAAGCAGTAGTGTAAACAGTCTTCCAGATGTCTCAGATGACAAGAT-----GAGCTAGAC 502
QY 65 AspileLeuSerThrTyrCysValAspAsnAsnGlnGly-----GlyProGly 80
Db 503 ATGATGAGGAGGAGACACACAGTGTCTTAACAGATTCTGTATCCATTTTAAACCCAGAG 562
QY 81 GluAspGly-----AlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArg 98
Db 563 GAAGAAAATACAGAGAGAGAGGAGATCTCT-----AGA 595
QY 99 ThrTyrValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGluLysGlu 118
Db 596 ACA-----AGAGCATCAGATCCCAATCTCCACCCCAAGTATCTCGTCACAAATCA 646
QY 119 ProSerLysGlyAspProAsnThrGluGluLeuArgGlnSerAspGluValGlyAspArg 138
Db 647 CACTATCGTATTCAGAGACACACTTTGCTACTATACGAGCAGCATCCTGGTT----- 697
QY 139 AspHisArgArgProGlnGlyLysLysLysLysLysLysLysLysLysLysLysLysLys 157
Db 698 ---ACGAGGCAAAATGCAAGAACATGAGCAGGACTCTGAGCTTAGAGAACAAATGCTGCG 754
QY 158 -----LeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAla 174
Db 755 TATAAGCGAATGAGCGGCAACATCAAAAGCAACTGATGACTCTGGAACCAACAGCTAAAG 814
QY 175 AlaLeuCysLysLysTyrAla-----GluLeuLeuGluGluHisArgAsn----- 189
Db 815 GCTGAGATGATGAACATCGCCTCAGATTAGAACAAAGATCTTGAATCAGCTAACAAAT 874
QY 190 ---SerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGlu----- 207
Db 875 TTTGCTGCAGAAATCGAGAACTTATCAAGAAACACACAGGCTGCCATGGAGAAAGAGGCT 934
QY 208 -----LysAspHisLeuArgGlyGluHisSerLys 217
Db 935 AAAGTGATGTCCAATGAAGAGAAAATTTTCAGCAACATATTCAGGCCCAACAGAGAAA 994
QY 218 AlaValLeuAlaArgSerLysLeuGluSerLysGluLeuGlnArgHisAsnArg 237
Db 995 GAACGTG-----AATAGTTTCTCGAGTCCAGAAAGAGAGATATAAATCTCGAAAAGAG 1048
QY 238 SerLeuLysGluGluGlyValGlnArg-----AlaArgGluGluGluGlu----- 252
Db 1049 CAGCTTAAAGAGGAGCTAAATCAAAACCCAGAGTAGTACCCCAAAAGAGAGAGAGAGAGTGG 1108
QY 253 -----LysArgLysGluValThrSerHisPheGlnVal-----ThrLeuAsnAspile 268
Db 1109 CTTTCAAGCAGAGGAGAGATATACAGCATTTTCCAGCAGAGAGAGAGAGAGAGAGAGAG 1168
QY 269 GlnLeuGlnMetGlu----- 273
Db 1169 CGAGCTCAAGAGACAATACCTAGAGCTGGAATGCCGCTTCAGAGAGAGAGAGAGAGAGAG 1228
QY 274 ---GlnHisAsn-----GluArgAsnSerLysLeuArgGln 284

```

Db 1229 GGGCGTCATCACTTAGACGAGGACCTTGTTCAGGAGGAGTAAACAAACACAGACTCAG 1288
Qy 285 GluAsnMetGluLeuAlaGluArgLeuLysLeuIleGluGlnTyrGluLeuArgGlu 304
Db 1289 AAGACTTAGACATGCGCTACTCTCGACAGCATGAATCTATGCAAGACTGGAGTTC 1348
Qy 305 GluHisIleAspLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLys 324
Db 1349 CGCCACCTCAACATTCAGAG-----ATGGCTGTGAGTTGATC-----AGA 1393
Qy 325 LeuGlnGlnAlaGlnGluMetLeuLysGluAlaGluLysGlnArgHisGlnArgGluLysAsp 344
Db 1394 TTACGATCAAACTGAGCTCTACCACTGCTGATATATTAAGCAGAGACGAGAA 1453
Qy 345 PheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlu 364
Db 1454 CTAAGACGAAAGCATGTCATGGAAGTTTCGACACAGCCTAAGAGTTTGAAGTCTAAGAA 1513
Qy 365 ThrHisLeuLysGlnGlnLeu-----AlaLeuTyrThrGluLysPheGluGlu 380
Db 1514 CTCGAAATATAAAGCAGTTTCAGATPACCTGCAAAATCCAAACACAGACAGTCAAGCA 1573
Qy 381 PheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGlu 400
Db 1574 TTAAGAAATCACTGCTGGAG-----ACTACCAAAAGAGTGAGCAAA 1618
Qy 401 LysMetThrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 420
Db 1619 GCTGTTCTGAACGGCTCAAG-----GAGGAACAGACCGGAAATAGCTATCTTGGCT 1672
Qy 421 GluSerSerAsnLysAlaLeuLeuGluMetAlaGluLysThrValArg----- 437
Db 1673 GACAGTATGATCACAGCATATGAATGCTCTCCACACAGCCTGCTTGGATGAA 1732
Qy 437 ----- 437
Db 1733 GCACAGGAGCAGAGTCCAGGTTTGAAGATGACGCTGCAGCAGCAACTGGAGCTGTG 1792
Qy 438 -----AspLysGluLeuGlu 442
Db 1793 AATGCGTATCAGACGAAATCAAGATGCAAGTGAAGTGAAGCAACATGATCAGAGCTTCG 1852
Qy 443 GlyLeuGlnValLysLys-----GlnArgLeuGluLysLeu 454
Db 1853 GAGCTTGAACAGAGGCTCTCCCTCGAGGCGACTCTTAGAACAAAGATTCAGAGAG 1912
Qy 455 CysArgAlaLeuGlnThrGluArgAsnAsp-----LeuAsnLysArgVal 469
Db 1913 ATGTTGGCTTTGAGAAATGAGCGCACAGCAACGAATACGAAGCCTGTTGGAACGTCAAGCC 1972
Qy 470 GlnAspLeuSerAlaGlyGlnGlyLys-----Leu 480
Db 1973 AGAGATTCAGACTTTGACTCTGAAGCATGAGACTAGGTTTAGTAATATGTCCTT 2032
Qy 481 ThrAspSerGlyProGluArgArgProGluGlyProGlyAlaGlnAlaProSerPro 500
Db 2033 TCTAATCTCTCCCTCAG----- 2050
Qy 501 ArgValThrGluAlaProCysTyrProGlyAlaPro-----SerThrGluAlaSerGly 518
Db 2051 -----GCATTACGACAGCTACCCGGAGCTTCTGGTTGTCACACACCTACTGG 2104
Qy 519 GlnThrGlyPro 522
Db 2105 GGTCCAGGACCT 2116

RESULT 35

US-09-023-655-1463
; Sequence 1463, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart

; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1463:
; LENGTH: 7596 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 9641957
; US-09-023-655-1463

Alignment Scores:
Pred. No.: 148e-11 Length: 7596
Score: 248.50 Matches: 144
Percent Similarity: 39.67% Conservative: 96
Best Local Similarity: 23.80% Mismatches: 226
Query Match: 9.20% Indels: 140
DB: 4 Gaps: 28

US-10-023-523-8 (1-530) x US-09-023-655-1463 (1-7596)

Qy 1 LysSerSerProGlyGlnProGluAla-----GlyProGluGlyAla--- 14
Db 4243 AAGTCTCGAAGAACCCCAAGAGAGCTTCTCAAGGACGCGGAGCCCTGTGACCGCGCCT 4302
Qy 15 -----GlnLysProSerGlnAlaProAlaValGlu 26
Db 4303 GGAGGAGAGCAGCTGCGCTATGCAAACTGGAGAGACCAAGACCGCTGCGAGCAGGA 4362
Qy 27 AlaGluGlyProGlySerSerGlnAlaProArgLysProGluGlyAlaGlnAlaArgThr 46
Db 4363 GCTGGACGACCTCAGCGTGGACCTGGACCAACCGCCAGCGCTGCGCTCCAA----- 4413
Qy 47 AlaGlnSerGlyAlaLeuArgAspValSerGluGlu-----LeuSerArgGln 62
Db 4414 -----CTTGGAGAGAGCAGAGAGATTTCACCACTGTAGCAGAGA 4458
Qy 63 LeuGluAspIleLeuSerThrTyrCysValAspAsnGlnGlyProGlyGluAsp 82
Db 4459 ---GAAGAGCATCTCTGCTCGCTATGC-----CGAAGAGCGGAGCCGCGCGCAAGC 4506
Qy 83 GlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyrValAla 102

Db 4507 CGAGGCCAGAGAGAA----- 4521
QY 103 ArgAsnGlyGluProGluProThrProValValTyrGlyGluLysGluProSerLysGly 122
Db 4522 AGAACCAAGCCCTGCTACTGGCCG-----GGCCCTCGAGGA 4560
QY 123 AspProAsnThrGlu---GluLeuArgGlnSerAspGluValGlyAspArgAspHisArg 141
Db 4561 AGCCCTGGAGGCCAGGAGGAGTTTGAAGAGGAGAGAACACAGCAGCTCCGAGCAGACATGGA 4620
QY 142 ArgProGlnGlu-LysLysLysAlaLysGlyLeuGlyLysGluLeuThrLeuLeuMetG1 161
Db 4621 AGACCTCATGAGCTCCAAA-----CATGATGTGGGAAAAACGTTCCAGCACTTGAATA 4674
QY 161 nThrLeuAsnThrLeuSerThrProGluGlnLysLeuAlaLeuCysLysLysTyrAl 181
Db 4675 ATCCAAAGCCCTAGACAGCAGGTGGAGGAATG-----AGGACCCAGCT 4722
QY 181 aGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLysLeu----- 196
Db 4723 GGAGGAGCTGGAGAGCAATCTCCAGGCCACGGAAGATGCCAAGCTTCGTCTGGAGGTCAA 4782
QY 197 -LeuGlnLysLysGlnSerGlnLeu-----ValGlnGluLysAspHisLeuArgG1 213
Db 4783 CATGCGAGCCATGAAGCGCAGTTCCGAGAGACCTCGCAACACAGGAGTGA---CAGAA 4839
QY 213 yGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuG1 233
Db 4840 TGAAGAGAAAGAGCGCTCTGATCAACAG-----GTCCGGGAGCTCGA 4884
QY 233 nArgHisAsnArgSerLysGluGluGluValGlnArgAlaArgGluGluGluGly 253
Db 4885 G-----GGGAGCTGAGATGAGTATTAGCAGCTCCGAGCTCCAGCTCAGAT 5040
QY 253 sArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuMetG1 273
Db 4936 GAAAAAG-----ATGAGATAGACCTTGAAGGAGCTCGAAGCCCAATCGA 4980
QY 273 uGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAsnMetGluLeuAlaGluLeu 293
Db 4981 GGTGCGCAACAAAGCTCGGATGAGTATTAGCAGCTCCGAGCTCCAGCTCAGAT 5040
QY 293 uLysLysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAspLysValPhe----- 311
Db 5041 GAAGGATTACCAACGTGAATTAGAAAGAGCTCGTGCATCCAGAGATGAGATTTGCTCA 5100
QY 312 ----LysHisLysAspLeuGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnG1 330
Db 5101 ATCCAAAGAGAGTGAAGAAATTTGAAGAGTCTTGAAGCAGAAATCTTCAATTCGAGA 5160
QY 330 uMetLeuLysGluAlaGlu-----GluArgHis---GlnArgGluLysAspPheLeuLe 347
Db 5161 GGAAGTTCCTCATCTGAGCGAGCCCGCCGACACGCGGAGCAGGAGAGATGAGCTGC 5220
QY 347 uLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLe 367
Db 5221 GGAAGAGATCAACCAACAGCCCTCTGGCAAGTCCGCGCTGCTGGATGAGAAGCGCGCT 5280
QY 367 uLysGlnGlnLeuAlaLeuTyrThrGluLysPheGluPheGlnAsnThrLeuSerLys 387
Db 5281 GGAAGCTCGATCGCAGCTGAGGAGGAGGAGCTGGAGAGGAGCAGCAGCACTGAGCT 5340
QY 387 sSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLe 407
Db 5341 GCTCAACGACCGCTTC-----CGCAAGACCACTCTACAGGTGA 5379
QY 407 sLysLeuGluLysGluThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAla-- 426
Db 5380 CACACTGAACCGCGAGCTAGCAGCCGAGCGAGCGCGCCGAGAGAGTGAACATGACG 5439
QY 427 -----LeuLeuGluMetAlaGluGluLysThrValArgAspLysGluLeuGluGly-- 443

Db 5440 CCAGCAACTGGAGCGGCAGAGAACAGAGCTGAAGGCCAAGCTGCAGGAACCTCGAGGGTGC 5499
QY 444 -----LeuGlnValLysIleGlnArgLeuG1 452
Db 5500 TGTCAAGCTTAAGTTCAAGGCCACCATCTCAGCCCTGAGGCCCAAGATTGGGAGCTGGA 5559
QY 452 uLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsn----- 466
Db 5560 GGAGCAGCTTGAGCAGAGAGCCAGGAAACGAGCAGCCGCCCAAAATTAGTCGCTCCGCAC 5619
QY 467 ----LysArgValGlnAspLeuSerAlaGlyGlnGlnLysLeuThrAspSerGlyPr 485
Db 5620 TGAGAGAAGAGCTGAAGAAATCTT-CATGACAGGTTGAGGATGAGCGTCGACGCGGACC 5678
QY 485 oGluArgArgProGluGlyProGlyAlaGln-----AlaPr 497
Db 5679 AGTATAAGAGCAGATGGAGAGGCCAACGCTCGGATGAAGCAGCTTAAAGCCAGCTGG 5738
QY 497 oSerSerProArgValThrGluAla-----Pro----- 506
Db 5739 AGGAAGCAGAGAGAGAGAGCGAGCGTCCCAACGCTCTCGGCGTAAACTCCAGCGGGAAC 5798
QY 507 -----CysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnG1 524
Db 5799 TGGATGATGCCACCGAGCGCAACGAGCGCTGAGCCGCGAGGTGAG---CACCCCTGAAGA 5855
QY 524 uProThrSerAla 528
Db 5856 ACCGCTGAGCGG 5868
RESULT 36
US-09-060-410-1
; Sequence 1, Application US/09060410
; Patent No. 6165461
; GENERAL INFORMATION:
; APPLICANT: Cobb, Melanie
; APPLICANT: Hutchinson, Michele
; APPLICANT: Chen, Zhu
; APPLICANT: Berman, Kevin
; TITLE OF INVENTION: TAO PROTEIN KINASES AND METHODS OF USE
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,410
; FILING DATE: 14-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 860098.421
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3312 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS

LOCATION: 121..3123
US-09-060-410-1

Alignment Scores:
Pred. No.: 5,89e-12 Length: 3312
Score: 247.00 Matches: 142
Percent Similarity: 40.12% Conservatives: 118
Best Local Similarity: 21.91% Mismatches: 212
Query Match: 9.14% Indels: 176
DB: 3 Gaps: 30

US-10-023-523-8 (1-530) x US-09-060-410-1 (1-3312)

QY 23 ProAlaValGluAlaGlu----- 28
DB 1090 CCAGCAGTAGAAGCAGGAGAGAGAGAGAGAGATCATGGTGGTGGCCGACAGGA 1149

QY 29 -----GlyProGlySerSerGlnAla---ProArgLysProGluGlyAlaGlnAla 44
DB 1150 ACAGTAATAGTGTGGAGCAATCAGTCTATCCAGATGCTATCATGTCAGTGCAGTAGC 1209

QY 45 ArgThrAlaGlnSerGlyAlaLeuA:GAspValSerGluLeuSerArgGlnLeuGlu 64
DB 1210 CAAAGCAGCAGTGTAAATGCTCTCCAGATGCTATCGATGCAAGAGT---GAGCTAGAC 1266

QY 65 AspileuSerThrTyrcysvalaspAsnAsnGlnGly-----GlyProGly 80
DB 1267 ATGATGAGGAGGAGCAGCAGTATGCTATTAACAGTCTGCTATCCACTTAAACCTGAG 1326

QY 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTy 100
DB 1327 GAGGAAATACCA---GAAGAGGAGATCCTAGA----- 1359

QY 101 ValAlaArgAsnGlyGluProGluProThrProValValTyrcysGluLysGluProSer 120
DB 1360 ---ACAAGCAGCAGCAGTCCACAGTCTCCACCTCAAGTGTCTGCTCACAATACATAT 1416

QY 121 LysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHis 140
DB 1417 CGTAATAGAGAACACTTTGCACTATACGACAGCATCAGTGT-----ACA 1464

QY 141 ArgArgProGlnGlyLysLysAlaLysGlyLeuGlyLysGluIleThr----- 157
DB 1465 AGACAGATGCAAGAACATGACGAGGACTCTGGAACCTTAGACAGACAGATGCTGGTTATAAG 1524

QY 158 ---LeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGlyLysLeuAlaLeu 176
DB 1525 CGGATGAGGCGCAGCATCAGAGCAGCTGATGACTCTGGAATAATAACTGAAGCAGAA 1584

QY 177 CysLysLysTyra-----GluLeuLeuGluGluHisArgAsn-----Ser 190
DB 1585 ATGCAGCAACATCGGCTCAGATTAGACAAAGATCTTGAACCTCAGCGCAACAATTCGCT 1644

QY 191 GlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGlu----- 207
DB 1645 GCAGAAATGAGAAACTATTAAAGAACACCAAGCTTCTATGGAAGAAAGAGGCTAAAGTG 1704

QY 208 -----LysAspHisLeuArgGlyGlyHisSerLysAlaVal 219
DB 1705 ATGCCCAACAGGAGAGAAATCCAAACACACATTCAGGCTCAACAGAGAAAGAACTG 1764

QY 220 LeuAlaArgSerLysLeuGluSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeu 239
DB 1765 -----AATAGCTTTTGGAGTCTCAAAAAGAGAAATATAAAATTCGAAAAGAGGCTAAAGTG 1818

QY 240 LysGluGlyValGlnArg-----AlaArgGluGluGlu----- 252
DB 1819 AAGGAGGAGCTGATGAAACACAGACACACCTTAAAGAAAGAAAGCAGGAATGGCTTCA 1878

QY 253 LysArgLysGluValThrSerHisGlnVal-----ThrLeuAsnAspLysGlnLeu 270
DB 1879 AAGCAGAGGAGGAATATTCAACATTTTCAGGAGAGAGAGAGTAAATCTCTCTCGACGT 1938

QY 271 GlnMetGlu-----Gln 274
DB 1939 CAAAGGCAGTATCTAGAGCTAGATGTCTGCTTCAAAAGAGAGATGTTACTTGGTCGG 1998

QY 275 HisAsn-----GluArgAsnSerLysLeuArgGlnGluAsn 286
DB 1999 CATACCTTGGAAACAGGACCTTGTCCAGGAGAGGATTAACAAAGAGCAGACTCAGAGAGAC 2058

QY 287 MetGluLeuAlaGluArgLeuLysLeuIleGluGlnTyrcysGluLeuArgGluGluHis 306
DB 2059 TTAGAACATGCAATGTTACTCCGACAGCATGATCCATGCAAGAACTGGAGTTCGCCAC 2118

QY 307 IleAspLysValPheLysHisLysAspLeuGlnGlnGlnLeuValAspAlaLysLeuGln 326
DB 2119 CTCACACTATTTCAGAAAG-----ATGCGCTGTGAGTTGATC-----AGACTCAA 2163

QY 327 GlnAlaGlnGluMetLeuLysGluAlaGluArgHisGlnArgGluLysAspPheLeu 346
DB 2164 CATCAACTGAGCTTACTTAACAGCTGGAATACATAAGAGAGGAAACGGAACTAAGA 2223

QY 347 LeuLysGluAlaValGluSerGlnArgMetCysGluLeuMetLysGlnGlnGlnThrHis 366
DB 2224 CGGAAACATGTCATGGAAGTTCGACAGCAGCTAAGAGTTTGAAGTCTAAAGAACTCCAA 2283

QY 367 LeuLysGlnGlnLeu-----AlaLeuTyrcysGluLysPheGluGluPheGln 382
DB 2284 ATAAAGAGCAGTTTCAGGATACCTGCAAAATTCAAACACAGACAGTACAAAGCATTAAAG 2343

QY 383 AsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMet 402
DB 2344 AATCACTTACTGGAG-----ACTACACCAAGAGTGAAGCACAAGCTGTT 2388

QY 403 ThrLysLysIleLysLysLeuGluLysGluThrThrMetTyrcysArgSerArgTrpGluSer 422
DB 2389 CTGAAAGACCTCAAG-----GAGAAACAGACTCGGAAGTTAGCCATCTTGGCTGAGCAG 2442

QY 423 SerAsnLysAlaLeuLeuGluMetAlaGluLysThrValArg----- 437
DB 2443 TATGATCATAGCATTAATGAATCTCTCCACACAGCTCTGCTGTTGGATGAAGCAGAC 2502

QY 437 ----- 437

DB 2503 GAAGCAGAAATCCAGGTTTGAAGATGACGCTACAGCAGGAACTGGAGCTGTGAATGCA 2562

QY 438 -----AspLysGluLeuGluGlyLeu 444
DB 2563 TATCAGACCAAAATCAAGATCCAGGCTGAGGCCCAACATGATCGAGAGCTTCGAGAGCTG 2622

QY 445 GlnValLysIle-----GlnArgLeuGluLysLeuCysArg 456
DB 2623 GAACAAAGGGTCTCCCTTCGGAGAGCAGCTCTTAGAACAGAGATTGAAGAGAGATGTTG 2682

QY 457 AlaLeuGlnThrGluArgAsnAsp-----LeuAsnLysArgValGlnAsp 471
DB 2683 GCTTTGCAAGATGAACGACAGAAAGATAGCTAGCTGCTCGAGCCCGAGCCAGAGAA 2742

QY 472 LeuSerAlaGlyGlyGlnGlySer-----LeuThrAsp 482
DB 2743 ATTGAAGCTTTTGAATCTGAAAGCATGAGATTAGTGTATTAACATGCTTCTTCTTAAT 2802

QY 483 SerGlyProGluArgArgProGluGly---ProGlyAlaGlnAlaProSer---SerPro 500
DB 2803 CTCTCCCTTGGAGCATTCAGCCACAGCTACCCAGAGACTTCTAGTGGTCTCAATCTCT 2862

QY 501 ArgValThrGluAlaPro-----CysTyrcysProGlyAlaProSerThrGluAlaSerGly 518
DB 2863 ACTGGGGTTTCAGGACCTCACTGGGGTCAATCCATCGGTGGCAGCACCACCAAGCTTGGGT 2922

QY 519 -----GlnThrGlyProGln 523
DB 2923 CATCCGATGCAAGCGGACCCCAA 2946

RESULT 37

US-09-723-458-1
 ; Sequence 1, Application US/09723458
 ; Patent No. 6586242
 ; GENERAL INFORMATION:
 ; APPLICANT: Cobb, Melanie
 ; Hutchinson, Michele
 ; Chen, Zhu
 ; Berman, Kevin
 ; TITLE OF INVENTION: TAO PROTEIN KINASES AND METHODS OF USE
 ; THEREFOR
 ; NUMBER OF SEQUENCES: 26
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SED and BERRY LLP
 ; STREET: 6300 Columbia Center, 701 Fifth Avenue
 ; CITY: Seattle
 ; STATE: Washington
 ; COUNTRY: USA
 ; ZIP: 98104
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09723458
 ; FILING DATE: 27-Nov. 6586242-2000
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/060,410
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Maki, David J.
 ; REGISTRATION NUMBER: 31,392
 ; REFERENCE/DOCKET NUMBER: 860098.421
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (206) 622-4900
 ; TELEFAX: (206) 682-6031
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3312 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 121..3123
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 ; US-09-723-458-1

Alignment Scores:
 Pred. No.: 5.89e-12 Length: 3312
 Score: 247.00 Matches: 142
 Percent Similarity: 40.12% Conservative: 118
 Best Local Similarity: 21.91% Mismatches: 212
 Query Match: 9.14% Indels: 176
 DB: 4 Gaps: 30

US-10-023-523-8 (1-530) x US-09-723-458-1 (1-3312)

1267 ATGATGGAGGAGGAGCATACAGTGTCTAAACAGTTCTGTCTCCACTTAAACCTGAG 1326
 81 GluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGluLysSerArgThrTyr 100
 1327 CAGGAAATATTACCAA--GAAGAAGGAGATCTCAGA----- 1359
 101 ValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGluLysGluProSer 120
 1360 ---ACAAGAGCATCAGCTCCACAGTCTCCACCTCAAGTGTCTCGTCACAATCATCATAT 1416
 121 LysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHis 140
 1417 CGTAATAGAGAACACTTTTGCACCTATATACGAACAGCATCATCTGGTT-----ACA 1464
 141 ArgArgProGlnGlnLysLysAlaLysGlyLysGlyLysGluIleThr----- 157
 1465 AGACATGCAAGAACACATGACGAGACTCTGAATTTAGAACACAGATGTCGTGTATTAAG 1524
 158 ---LeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeu 176
 1525 CGGATGAGCGGACAGCATCAGACGAGCTGATGACTCTGGAAATAAATCACTGAAGCGAGAA 1584
 177 CysLysLysTyrAla-----GluLeuLeuGluGluHisArgAsn-----Ser 190
 1585 ATGGACGAACATCGGCTCAGATTAGACAAAGATCTTGAACTCAGCGCAACAATTCGCT 1644
 191 GlnLysGlnMetLysLeuGlnLysGlnSerGlnLeuValGlnGlu----- 207
 1645 GCAGAAATGGAGAACTATTAAAGAACACCAAGCTTCTATGGAAGAAAGAGCTTAAAGTG 1704
 208 -----LysAspHisLeuArgGlyGlyGluHisSerLysAlaVal 219
 1705 ATGGCAACAGGAGGAAAAAATTCACACACATCTCAGGCTCAACAGAGAAAGAACTG 1764
 220 LeuAlaArgSerLysLeuGluSerLysCysArgGluLeuGlnArgHisAsnArgSerLeu 239
 1765 -----AATAGCTTTTGGAGTCTCAAAAGAGAGATATAAACTTCGAAAAGAGAGCTT 1818
 240 LysGluGluGlyValGlnArg-----AlaArgGluGluGluGlu----- 252
 1819 AAGGAGGAGCTGAATGAAACACAGAGCACACCTAAAAAGAAAGAGCAGGAATGGCTTCA 1878
 253 LysArgLysGluValThrSerHisPheGlnVal-----ThrLeuAsnAspIleGlnLeu 270
 1879 AAGCAGAAGAGAGATATTCACATTTTCAGGCGAGAGAGAGTAACTCTTCTTCGAGCT 1938
 271 GlnMetGlu-----Gln 274
 1939 CAAAGCGAGTATCTAGAGCTAGATGTCTCGCTTCAAAAGAGAAATGTTACTTGTGCGG 1998
 275 HisAsn-----GluArgAsnSerLysLeuArgGlnGluAsn 286
 1999 CATAACTTGGAAACAGGACCTTGTTCAGGAGGAGGTAAACAAAGAGCAGACTCAGAGGAC 2058
 287 MetGluLeuAlaGluArgLysLysLeuIleGluGlnTyrGluLeuArgGluGluHis 306
 2059 TTAGAACATGCATGTTACTCGCAGCATGATCCATCCATGCAAGACTGGAGTTTCGCCAC 2118
 307 IleAspLysValPheLysHisLysAspLeuGlnGlnLeuValAspAlaLysLeuGln 326
 2119 CTCACACATATTCAAGAG-----ATGGCTGTGAGTTGATC-----AGACTGCAA 2163
 327 GlnAlaGlnGluMetLysGluAlaGluArgHisGlnArgGluLysAspPheLeu 346
 2164 CATCAACTGAGCTTACTTAACCGCTGGAAATACATAAGAGAGGAAACGGGAACATAAGA 2223
 347 LeuLysGluAlaValGlnSerGlnArgMetCysGluLeuMetLysGlnGlnThrHis 366
 2224 CGGAAACATGTCATGTAAGTTCGACAGCAGCTTAAGAGCTTGAAGTCTTAAAGAACTCAA 2283
 367 LeuLysGlnGlnLeu-----AlaLeuTyrThrGluLysPheGluGluPheGln 382
 2284 ATAAAAAGAGAGTTTCAGGATACCTGCAAAATTCACACCAGACAGTACAAAGCATTAAGG 2343


```
QY 368 LysGlnGlnLeuAlaLeuTyThrGluLysPheGluGluPheGlnAsnThr-----Leu 385
Db 3451 AAACCTACAGATGAATTTGGACACACTTAAAGAAAAACAACTTTGAAAAATGTGGAAGAGCTG 3510
QY 386 SerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLys 405
Db 3511 AACAAATCAAAGAACTCCTGACTGTAGAGAAATCAAAATATGGAGAAATTTAGGAAGAA 3570
QY 406 IleLysLysLeuGluLysGluThrThr-----414
Db 3571 ATAGAAACCTTAACGACGACGACGCTCAGAAGTCCGACGAGCTTTCAGCGTTGCAAGAA 3630
QY 415 -----MetTyrArgSerArgTyrPheGluSerSerAsnLysAla 426
Db 3631 GAGAAAGCTTAAACTTCCTGAGGAGCTGGGAGACGCGGACGAGTCAACAAGTCATCAA 3690
QY 427 LeuLeuGluMetAlaGluGluLysThrValArgAspLysGluLeuGluGluLysVal 446
Db 3691 AAGCTGGAA-----GAGAAAGATCTGTGCTCAATAATCAGTTG-----TTAGAAATG 3738
QY 447 LysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsn 466
Db 3739 AAA-----AAAAGAGATCCAAAGTTTCATAAAGACGACGAGATGAGAGAAAGCTTCCTTGCA 3795
QY 467 LysArgValGlnAspLysSerAla 474
Db 3796 AAATCCATCAGTAACTAGTGCC 3819

RESULT 39
US-09-976-594-268.
; Sequence 268, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 268
; LENGTH: 6617
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 994387.65
US-09-976-594-268

Alignment Scores:
Pred. No.: 2,71e-11 Length: 6617
Score: 244.50 Matches: 143
Percent Similarity: 40.78% Conservative: 107
Best Local Similarity: 23.33% Mismatches: 177
Query Match: 9.05% Indels: 189
Db: 4 Gaps: 30

US-10-023-523-8 (1-530) x US-09-976-594-268 (1-6617)
QY 2 SerSerProGlyGlnProGluAlaGly-----ProGluGlyAla-----14
Db 807 GCAGCAGCACCACGACGAGTAGCTGGAGCCCGACGCTCCCAAGAGAGCTTGGATGTCGGA 866
QY 15 -----GlnGluArgProSerGlnAlaAlaProAlaVal-----25
Db 867 CAAGCAGGGTGGACTCCCGAGGACGACCCAGGCTTGGTGGAGAGAGCTGTGTGTA CAGA 926
QY 26 GluAlaGluGlyProGlySerSerGlnAlaPro-----ArgLysProGluGly 41
Db 927 GAGCGGCAGGA-----GGAGCAGGAGAGGCTGCATGCACTTTTCGGGCTGGGAGAAC 983
```

```
QY 41 -----41
Db 984 AGGAGCTAAACATCACCATCATCTAAGAGCGGTCAAGAAATTGAAAAAACAACA 1043
QY 42 -----AlaGln 43
Db 1044 ACATTTAAGGGTTAATATCTTACACAAATTCATTACTTCAATTGATGTTAGAGCCCC 1103
QY 44 AlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArgGlnLeu 63.
Db 1104 AGCAGGCCATCTTAAGGTAGAGCAGCTACAAGATGAGACCAACACCCTAAGGAAGGAGCTA 1163
QY 64 GluAspIleLeuSerThrTyrCysValAspAsnGlnGlyGlyProGlyGluAspGly 83
Db 1164 GAGAGTGGG-AAGACA-----CCTCAGGCTGAGGTGA 1198
QY 84 AlaGlnGlyProAlaGluProGluAspAlaGluLysSerArgThrTyrValAlaArg 103
Db 1199 AAACAATCATGATGTGAGTCTCTCTGAACA-GGAGACAGGAGGAGGCTACGTGAACAGG 1257
QY 104 AsnGlyGluProGluProThrProValValTyrGlyLysGluProSerLysGlyAsp 123
Db 1258 AGG-----AGAGGCTAC 1269
QY 124 ProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHisArgArgPro 143
Db 1270 GTGAACAG-GAGGAGAGGCAACGCTGAACAGGAG-----GATAGGCTACATGAACAGGAG 1322
QY 144 GlnGlu-----LysLysLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThr 162
Db 1323 GAGAGGCTACGTGAACAGGAGGAGGCTGTGTGAACAGGAGGAGGCTGTGTGAACAG 1382
QY 163 LeuAsnThrLeuSerThrProGluGluLysLeuAla-----AlaLeuCysLys 178
Db 1383 GAGAGAGGCTACGTGAACATGAGGAGGCTGTGTGAACAGGAGGAGGCTATGTGA 1442
QY 179 LysTyrAlaGluLeuLeuGluHisArgAsnSerGlnLysGlnMetLysLeuGln 198
Db 1443 CAGGAGGAGGCTACGTGAACAGGAGGAGGCTACATGAACAGGAGGAGGCTACGT 1502
QY 199 LysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHisSerLysAla 218
Db 1503 GAACAGGAGGAGGCTGTGTGAACAGGAGGAGGCTACGT-----GAACATGAG-----1553
QY 219 ValLeuAlaArgSerLysLeuGluSerLysLeuGluSerGluGlnArgHisAsnArgSer 238
Db 1554 -----GAGAGGCTGTGTGAACAGGAGGAGGAG-----1580
QY 239 LeuLysGluGluGlyValGlnArgAlaArgGluGluGluLysArgLysGluValThr 258
Db 1581 CTACGTGAACAG-----GAGGAGAGGCTGTGTGAACAGGAGGAGGCTACGTGA-----1631
QY 259 SerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGlnHisAsnGluArg 278
Db 1632 -----CAGGAGGAGGAGGCTGTGTGAACAG 1655
QY 279 AsnSerLysLeuArgGlnGluAsnMetGluLeuAlaGluArgLeuLysLysLeuGlu 298
Db 1656 GAGGAGAGGCTACGTGAACAGGAGGAGGCTGTGTGTGAACAG-----GAGAAGCTCCAGGG 1712
QY 299 GlnTyrGluLeuArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGln 318
Db 1713 CAGGAGAGGCTGTGTGAACAG-----GTGAGAGGCTGTGTGAACAGGAGGAGGCGCAGGAG 1769
QY 319 GlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGlu-----336
Db 1770 GAG-----CAGGAGAGGCTGTGTGAACAGGAGGAGGCTGTGTGAACAGGAGGCTG 1823
QY 337 -----GluArgHisGlnArgLysAspPheLeuLeuLysGluAlaValGluSer 353
Db 1824 TTAGAACAGGAGGAGGCTGTGTGAACAGGAGGAGGCTGTGTGTGAACAGGAGGCTGCTG-----GAGAGG, 1871
```

```
QY 354 GlnArgMetCysGlu-----LeuMetLysGlnGlnGlnThrHisLeuLysGln 369
Db 1872 GAGAGGCTGCTGGAGAGGCTGGAAGAGCTGTTAGAACAGGAGAGCGCGAGGAGGAGCAG 1931
QY 370 GlnLeuAlaLeuThrThrGluLys---PheGluGluPheGlnAsnThrLeuSerLysSer 388
Db 1932 GAGAGGCTGCTGGAGAGGAGGCTGCTGCTGAGAGGCTGAGAGGCTCTGAGAGGAGTG 1991
QY 389 SerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysLysLys 408
Db 1992 GAGAGGCTCTG-----GAGCAGAGAGGCTTCGGCAACAGGATGAGAGG 2036
QY 409 LeuGluLysGluThrThrMetTyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeu 428
Db 2037 CTGGGAGGAGGAGACTCTG----- 2057
QY 429 GluMetAlaGluGluLysThrValArgAspLysGluGluGluGluValLysLys 448
Db 2058 -----CAGGAGCTGGAGAGGCTGGG---GAGCTG 2084
QY 449 GlnArgLeuGluLysLeuCysArgAlaLeuGln----- 459
Db 2085 GAGAGGCTGGGAGGCTGGAGAGGAGTGTGAGCTGGGCTGGGAGGAGGCTGACGAGCAG 2144
QY 460 -----ThrGluArgAsnAspLeuAsnLysArgValGlnAsp 471
Db 2145 CGGCGCAGCCAGCAGCGGCTTCGAGGAGCTCAACACGAGAGCAAGACACACTGCGAG 2204
QY 472 LeuSerAlaGluGlnGlnGlnSerLeuThrAspSerGlyProGluArgArgProGluGly 491
Db 2205 TTGAGCAGCAAGTAAG---GAGCTGAAGAGTCCGGTGGAGCTGAAGAGCCAGAGGC 2261
QY 492 ProGluValAlaGlnAlaProSerSerProArgValThrGluAlaProCysTyrProGluAla 511
Db 2262 TCCAGTCTGCAGCAGCAGCAGCAGCA---GTAGCTGGAGCCCCAGTCCCAAGGAGCT 2318
QY 512 ProSerThrGluAlaSerGlyGlnThrGlyProGlnGlu 524
Db 2319 TGGATGTGGCGGACAGCAGGCTGAGT---CCCCAGGAG 2354
```

RESULT 40

```
US-09-620-312D-248
; Sequence 248, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Felyan
; APPLICANT: Chen, Rui-Rong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aigong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yundong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP28
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
```

```
; SEQ ID NO 248
; LENGTH: 7453
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(4362)
US-09-620-312D-248

Alignment Scores:
Pred. No.: 5,94e-11 Length: 7453
Score: 241.50 Matches: 145
Percent Similarity: 38.45% Conservative: 98
Best Local Similarity: 22.94% Mismatches: 227
Query Match: 8.94% Indels: 165
DB: 4 Gaps: 26

US-10-023-523-8 (1-530) x US-09-620-312D-248 (1-7453)

QY 1 LysSerSerProGlyGlnProGluAlaGlyPro----- 11
Db 4074 CAGGAAGAACCTGGAGAAGCAAGTCTGGCCCTGCAGTCCAGTTGGCTGATACCAAGAA 4133
QY 12 -----GluGlyAlaGln 15
Db 4134 GAAAGTAGATGACGACCTGGGAACAATTGAAAGTCTGGAAGAAGCCAAAGAGAGCTTCT 4193
QY 16 GluArgProSerGlnAlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAla 35
Db 4194 GAAGGACCGGAGGCGCTGAGCCGCTGGAGGAGGAGGCACTGGCGTATGACAAACT 4253
QY 36 ProArgLysProGluGlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgSpVal 55
Db 4254 GG-AGAAGACCA-----AGAACCGCCTGCAGCAG-GAGCTGGAGCAGCTC 4296
QY 56 SerGluGluLeuSerArgGln----- 62
Db 4297 ACGGTGGACCTGGACCAACCGCCAGCGCTGCCCTCCAACTTGGAGAAGCAAGCAAGAA 4356
QY 63 -----LeuGluAspIleLeuSerThrTyrCysValArgAsn 75
Db 4357 GTTGACCACTGTAGCAGAAGAGAGCAGCATCTGCTCGCTATGC-----CGA 4407
QY 76 GlnGlyGlyProGlyGluAspGlyAlaGlnGlyGluProAlaGluProGluAspAlaGlu 95
Db 4408 AGACGGGACCGGGCCGAGCGGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4443
QY 96 LysSerArgThrTyrValAlaArgAsnGlyGluProGluProThrProValValTyrGly 115
Db 4444 -----AGAAACCAAGCCCTGCTCAGTGGCCCG----- 4470
QY 116 GlutylsGluProSerLysGlyAspProAsnThrGlu---GluIleArgGlnSerAspGlu 134
Db 4471 -----GGCCCTCGAGGAGCCCTGGAGGCGCAGGAGAGCTTTGAGAGGAGAGAGAG 4521
QY 135 ValGlyAspArgAspHisArgArgProGlnGlu-LysLysLysAlaLysGlyLeuGly 154
Db 4522 GCAGCTCCGAGCAGACATGGAAGACCTCATGAGCTCCAA-----GATGATGTGGGAAA 4575
QY 154 sGluIleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAl 174
Db 4576 AAACGTTCAAGACTTGAATAATCCAAACGGGCCCTAGAGCAGCAGCTGGAGAGAGATG-- 4633
QY 174 AlaLeuCysLysLysTyrAlaGluLeuLeuGluHisArgAsnSerGlnLysGlnMe 194
Db 4634 -----AGGACCCAGCTGGAGAGCTGGAAGAGCAACTCCAGCCACGAGATGC 4683
QY 194 tLysLeu-----LeuGlnLysLysGlnSerGlnLeu-ValG 206
Db 4684 CAAGCTTCGCTCGAGGTCAACATGAGGCCATGAAGGCCAGCTTCGAGAGAGACCTGCA 4743
QY 206 nGluLysAspHisLeuArgGlyLeuHisSerLysAlaValLeuAlaArgSerLysLeuGl 226
```



```
Db 4744 AACAGGATGAG---CAGATGAGAGAGAGAGCGGCTGCTGATCAACAG----- 4792
Qy 226 uSerLeuCysArgGluLeuGlnArgHisAsnArgSerLeuLysGluGluValGlnAr 246
Db 4793 -----GTGCGGAGCTCGAG-----GCCGAGCTGGAGGATCAGAGAAACAGCG 4836
Qy 246 gAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAs 266
Db 4837 GCGCTTCTAGCTTCAAGAAAGAG-----ATGAGATAGACCTTGAA 4881
Qy 266 nAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGluAs 286
Db 4882 GGACCTCGAAGCCCAATCGAGGCTGCGAACAAAGCTCGGATCAGGTATGACGAGCT 4941
Qy 286 nMetGluLeuAlaGluArgLeuLysLysLeuLysGlnThrGlyLeuArgGluGluHi 306
Db 4942 CCGCAAGCTCAGGCTCAGATGAGGATTAACCAACGCTGAATTAGAAGAAGCTCGTCATC 5001
Qy 306 sIleAspLysValPhe-----LysHisLysAspLeuGlnGlnLeuValAspAl 323
Db 5002 CAGAGATGAGATTTTGTCTCAATCCAAAGAGAGTGAAAGAAATGAGAGTCTGGAGC 5061
Qy 323 aLysLeuGlnGlnAlaGlnMetLeuLysGluAlaGlu-----GluArgHis----GI 340
Db 5062 AGAAATCTCAATTCAGGAGGAACTTGCTCATCTGAGCGAGCGCGCCGACAGCGCGA 5121
Qy 340 nArgGluLysAspPheLeuLysGluAlaValGluSerClnArgMetCysGluLeuMe 360
Db 5122 GCAGAGAGATGAGCTGGCGAGAGATCACCACAGCGCTCTGGCAAGTCGCGCT 5181
Qy 360 tLysGlnGlnGlnThrHisLysLysGlnGlnLeuAlaLeuThrThrGluLysPheGluG 380
Db 5182 GCTGGATGAGAAGCGGCGTCTGGAAGCTCGGATCGCACAGCTGAGGAGGAGCTGGAAGA 5241
Qy 380 uPheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPhelysGlnGluMetG 400
Db 5242 GGACGAGAGCAACATGAGCTGCTCAACGACCGCTTC-----CG 5280
Qy 400 uLysMetThrLysLysLysLysLysLysGluThrThrMetThrArgSerArgTr 420
Db 5281 CAAGACCACTCTACAGCTGGACACTGAACCGCGAGCTAGCAGCGAGCGAGCGCGCG 5340
Qy 420 pGluSerSerAsnLysAla-----LeuLeuGluMetAlaGluLysThrValAr 437
Db 5341 CCAGAGAGTGACAAATCCAGCCAGCAACTGGAGCGCGCAGAACAGAGGAGCTGAAGGCCAA 5400
Qy 437 gAspLysGluLeuGluGly-----LeuGlu 445
Db 5401 GCTCAGGAACTCAGGAGTGTCTCAAGTCTAAGTTCAGGCGCACCACTCTCAGCCCTGGA 5460
Qy 445 nValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLe 465
Db 5461 GGCCAGATTTGGCAGCTGGAGGAGCAGCTTGAGCAGGAAGCCAAAGAGCAGCAGCGCG 5520
Qy 465 uAsn-----LysArgValGlnAspLeuSerAlaGlyGlyGlnG 478
Db 5521 CAACAATATAGTCGCTCGACTGAGAAGAGCTGAAAGAAATCTT-CATGAGGTTGAGG 5579
Qy 478 ySerLeuThrAspSerGlyProGluArgArgProGluGlyProGlyValaGln----- 495
Db 5580 ATGAGCGTCGACACGCGGACCACTATAAAGAGCAGATGGAGAGGCCAACGCTCGGATGA 5639
Qy 496 -----AlaProSerSerProArgValThrGluAla----- 505
Db 5640 AGCAGCTTAAACGCCAGCTCGAGGAAGCAGAGAAGAAGCGCGCTGCCAACGCATCTC 5699
Qy 506 -----Pro-----CysTyrProGlyAlaProSerThrGluAlaSe 517
Db 5700 GGCGTAAACTCCAGCGGAACTGGATGATGCCACCGAGGCCAACAGAGGCGCTGAGCGCG 5759
Qy 517 rGlyGlnThrGlyProGlnGluProThrSerAla 528
Db 5760 AGGTCAAG---CACCTTGAAGAACCGGCTGAGGCG 5790
```

RESULT 41

```
US-09-620-312D-249
: Sequence 249, Application US/09620312D
: Patent No. 6569662
: GENERAL INFORMATION:
: APPLICANT: Tang, Y. Tom
: APPLICANT: Liu, Chenghua
: APPLICANT: Asundi, Vinod
: APPLICANT: Zhang, Jie
: APPLICANT: Ren, Feiyun
: APPLICANT: Chen, Rui-hong
: APPLICANT: Zhao, Qing A.
: APPLICANT: Wehrman, Tom
: APPLICANT: Xue, Aidong J.
: APPLICANT: Wang, Yonghong
: APPLICANT: Wang, Jian-Rui
: APPLICANT: Zhou, Ping
: APPLICANT: Ma, Yunging
: APPLICANT: Wang, Dunrui
: APPLICANT: Wang, Zhiwei
: APPLICANT: John Tillinghast
: APPLICANT: Drmanac, Radoje T.
: TITLE OF INVENTION: No. 6569662el Nucleic Acids and
: TITLE OF INVENTION: Polypeptides
: FILE REFERENCE: 784CIP28
: CURRENT APPLICATION NUMBER: US/09/620,312D
: CURRENT FILING DATE: 2000-07-19
: PRIOR APPLICATION NUMBER: 09/552,317
: PRIOR FILING DATE: 2000-04-25
: PRIOR APPLICATION NUMBER: 09/488,725
: PRIOR FILING DATE: 2000-01-21
: NUMBER OF SEQ ID NOS: 1105
: SOFTWARE: pt-FL_genes Version 1.0
: SEQ ID NO 249
: LENGTH: 7501
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)..(4410)
US-09-620-312D-249
```

```
Alignment Scores:
Pred. No.: 6e-11 Length: 7501
Score: 241.50 Matches: 145
Percent Similarity: 38.4% Conservative: 98
Best Local Similarity: 22.94% Mismatches: 227
Query Match: 8.94% Indels: 165
DB: 4 Gaps: 26
```

US-10-023-523-8 (1-530) x US-09-620-312D-249 (1-7501)

```
Qy 1 LysSerSerProGlyGlnProGluAlaGlyPro----- 11
Db 4122 CAGGAAGAACCTGGAGAGCAAGTCTGCGCTCGAGTCCAGTTCGCTGATACCAAGAA 4181
Qy 12 -----GluGlyAlaGln 15
Db 4182 GAAAGTAGATGACGACCTGGGACAAATTGAAAGTCTGGAAGAGCCAGAGAGCTTCT 4241
Qy 16 GluArgProSerGlnAlaAlaProAlaValGluAlaGluGlyProGlySerSerGlnAla 35
Db 4242 GAAGGACGCGGAGGCCCTGAGCCGCTGGAGGAGAGGACGCTGGCGGTATGACAAACT 4301
Qy 36 ProArgLysProGluGlyAlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspVal 55
Db 4302 GG-AGAGACCA-----AGAACCGCTGCAGCAG-GAGCTGAGCAGCTC 4344
Qy 56 SerGluGluLeuSerArgGln----- 62
Db 4345 ACGGTGAGCTTGGACCAACAGCGCGAGTCTCCACTTGGAGAGAGCAAGCAAGAA 4404
```

QY 63 -----LeuGluAspIleLeuSerThrTyrCysValAspAsn 75
 Db 4405 GTTTGACCAGCTGTAGCAGAGAGAGCATCTCTCGCTATGC-----CGA 4455
 QY 76 GlnGlyGlyProGlyGluAspGlyValGlnGlyProAlaGluProGluAspAlaGlu 95
 Db 4456 AGACGGGAGCCGCGGAGCCGAGCCGAGAGAA-----4491
 QY 96 LysSerArgThrTyrValAlaArgAsnGlyGluProGluProThrProValValTyrGly 115
 Db 4492 -----AGAAACCAAGCCCTGTCTAGTGGCCG-----4518
 QY 116 GluLysGluProSerLysGlyAspProAsnThrGlu-----GluIleArgGlnSerAspGlu 134
 Db 4519 -----GGCCCTCGAGGAGCCCTGGAGCCAGGAGAGTTTGGAGGGCGAGAA 4569
 QY 135 ValGlyAspArgAspHisArgProGlnGlu-LysLysLysAlaLysGlyLeuGlyLy 154
 Db 4570 GCAGCTCCGAGCAGACATGGAGACCTCATGAGCTCCAA-----GATGATGGGAAA 4623
 QY 154 sGluIleThrLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAl 174
 Db 4624 AAACGTTCCAGAACTTGAAATAATCCAAACGGGCGCTAGACAGCAGGTGGAGGAAATG-- 4681
 QY 174 aAlaLeuCysLysLysTyrAlaGluLeuGluGluHisArgAsnSerGlnLysGlnMe 194
 Db 4682 -----AGGACCCAGCTGGAGAGCTGGAGAGCGAACTCCAGGCCAGGAGATGC 4731
 QY 194 tLysLeu-----LeuGlnLysGlnSerGlnLeu-----ValGl 206
 Db 4732 CAAAGTTCTGCTGGAGGTCACATCCAGGCCATGAAGCGCAGTTCGAGAGACCTGCA 4791
 QY 206 nGluLysAspHisLeuArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGl 226
 Db 4792 AACCAAGGATGAG-----CAGAATGAAGAGAGCGCTGCTGATCAACAG-----4840
 QY 226 uSerLeuCysArgGluLeuGlnArgHisAsnArgSerLysLysGluGluGlyValGlnAr 246
 Db 4841 -----GTGCGGAGCTCGAG-----CGCGAGCTGGAGGATGAGAGAAACAGCG 4884
 QY 246 gAlaArgGluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAs 266
 Db 4885 GCGCTGCTGCTGAGCTTCAAGAAAAAG-----ATGGAGATGAGACCTGAA 4929
 QY 266 nAspIleGlnLeuGlnMetGluGlnHisAsnGluArgAsnSerLysLysGluGlnGluAs 286
 Db 4930 GGACCTCGAGCCCAATCGAGCTGCGAACAAAGCTCGGATGAGGTGATTAGCAGCT 4989
 QY 286 nMetGluLeuAlaGluArgLysLysLeuIleGluGlnTyrGluLeuArgGluGluHi 306
 Db 4990 CCGCAAGCTCCAGGCTCAGATGAGGATATCAACGCTGAATAGAGAAAGCTCGTGCATC 5049
 QY 306 sIleAspLysValPhe-----LysHisLysAspLeuGlnGlnLeuValAspAl 323
 Db 5050 CAGAGATGAGATTTTGTCAATCCAAAGAGTGAAGAAATGAAAGATGAGTCTGGAAGC 5109
 QY 323 aLysLeuGlnGlnAlaGlnGluMetLysGluAlaGlu-----GluArgHis---Gl 340
 Db 5110 AGAAATCCTTCAATTCGAGGAGGAACTTGCCTCATCTGAGCGAGCCGCGCAGACCGCGA 5169
 QY 340 nArgGluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeuMe 360
 Db 5170 GCAGGAGAGATGAGCTGGCGGAGAGATACCAACAGCGCCTCTGCAAGTCCGCGCT 5229
 QY 360 tLysGlnGlnGlnThrHisLysGlnGlnAlaLeuTyrThrGluLysPheGluGl 380
 Db 5230 GCTGGATGAGAGCGCGCTGGAAGCTCGATCGACAGCTGGAGGAGGAGCTGGAAGA 5289
 QY 380 uPheGlnAsnThrLeuSerLysSerGluValPheThrThrPheLysGlnGluMetGl 400
 Db 5290 GGAGCAGACCAACATGGAGCTGCTCAACAGCCGCTTC-----CG 5328
 QY 400 uLysMetThrLysLysLysLysGluThrThrMetTyrArgSerArgTr 420

Db 5329 CAAGACCACTCTACAGTGGACACTGAACGCCGAGCTAGCAGCGAGCGCGCG 5388
 QY 420 pGluSerSerAsnLysAla-----LeuLeuGluMetAlaGluGluLysThrValAr 437
 Db 5389 CCAGAAGAGTGAATGACACGCCCAACTGAGCGGAGGAGGAGCTGAAGGCCAA 5448
 QY 437 gaspLysGluLeuGluGly-----LeuGl 445
 Db 5449 GTTGAGGAAGTCTGAGGGTGTCTCAAGTCTAAGTTCAAGGCCCACTCTCAGCCCTGA 5508
 QY 445 nValLysIleGlnArgLeuGluLysCysArgAlaLeuGlnThrGluArgAsnAspLe 465
 Db 5509 GCGCAAGATTGGCAGCTGGAGGAGCAGCTTGAGCAGGAAGCAAGCAAGCAGCGCG 5568
 QY 465 uAsn-----LysArgValGlnAspLeuSerAlaGlyGlyGlnGl 478
 Db 5569 CAACAAATTAGTCCCTGCGACCTGAGAGAAGAGCTGAAGAAATCTT-CATGAGGTTGAGG 5627
 QY 478 ySerLeuThrAspSerGlyProGluArgProGluGlyProGlyValGln-----495
 Db 5628 ATGAGCTGACACCGGACCACTATAAGAGCAGATGAGAGAAGCCCAAGCTCGATGA 5687
 QY 496 -----AlaProSerSerProArgValThrGluAla-----505
 Db 5688 AGCAGCTTAACGCCAGCTGGAGGAGCAGAGAAGCAAGACGCGCTGCCAACGCATCTC 5747
 QY 506 -----Pro-----CysTyrProGlyAlaProSerThrGluAlase 517
 Db 5748 GCGGTAACTCCAGCGGAACTGGATGATGCCACGAGGCCCAACGAGGCGCTGAGCCGCG 5807
 QY 517 rGlyGlnThrGlyProGlnGluProThrSerAla 528
 Db 5808 AGTTCAG-----CAACCTGAAAGAACCGCTGAGGCG 5838
 RESULT 42
 US-09-595-684B-30
 ; Sequence 30, Application US/09595684B
 ; Patent No. 6544766
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; APPLICANT: Ohashi, Cara
 ; APPLICANT: Sakowicz, Roman
 ; APPLICANT: Vaisberg, Eugeni
 ; APPLICANT: Wood, Kenneth
 ; APPLICANT: Yu, Ming
 ; TITLE OF INVENTION: Human kinesins and methods of producing
 ; TITLE OF INVENTION: and purifying human kinesins
 ; FILE REFERENCE: cytop036
 ; CURRENT APPLICATION NUMBER: US/09/595,684B
 ; CURRENT FILING DATE: 2002-06-24
 ; PRIOR APPLICATION NUMBER: 09/295,612
 ; PRIOR FILING DATE: 2000-04-20
 ; NUMBER OF SEQ ID NOS: 105
 ; SOFTWARE: Fast-Seq for Windows Version 4.0
 ; SEQ ID NO 30
 ; LENGTH: 8257
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-595-684B-30
 Alignment Scores:
 Pred. No.: 9,36e-11 Length: 8257
 Score: 240.00 Matches: 130
 Percent Similarity: 35.81% Conservative: 101
 Best Local Similarity: 20.16% Mismatches: 176
 Query Match: 8.88% Indels: 238
 DB: 4 Gaps: 25
 US-10-023-523-8 (1-530) x US-09-595-684B-30 (1-8257)
 QY 55 valSerGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsn 74

Db 3049 ATTCTGAGGAAGTTTCCAGGAATTTGCGAT-----ATGAGGAA 3087
Qy 75 AsnGlnGlyProGlyGluAspGlyAlaGlnGly-----GluProAlaGlu 90
Db 3088 AATACAGGAAACTAAAGATGAATTTTCAGAAAGATGTTGGCATAGATAAAACAG 3147
Qy 91 ProGluAspAlaGluIysSerArgThrTyrValAla----- 102
Db 3148 GATTGGAAGCTAAATAACCAACACTAAGTGCAGATGTTAAGGATAATGAGATAAT 3207
Qy 103 -----ArgAsnGlyGluProGluProThr 110
Db 3208 GAGCAACAAAGGAGATATTTCTTTAATACAGGAGAAATGAACCTCCAAATGTTA 3267
Qy 111 ProValValTyrGlyGluGlyGluProSerLysGlyAsp----- 123
Db 3268 GAGAGTGTATAGCAGAAAGGAACTAATGAAGACTGACCTAAGGAAATATTGAAATG 3327
Qy 124 -----ProAsnThrGluGluIleArg-----GlnSerAspGluValGlyAspArgAspHis 140
Db 3328 ACCATTGAAACACAGGAGAAATTAAGACTCTCTGGGAGTGAATTTAAAGACACAAAG 3387
Qy 141 ArgArgProGlnGluIysLysAla-----LysGlyLeuGlyGlyGluIleThrLeuLeu 159
Db 3388 ATAGTTCACAGAAAGAACCATGCCATAAAGAAAGAGAGAG-----CTT 3435
Qy 160 MetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeu----- 173
Db 3436 TCTAGGACCTGTGACAGACTGGCAGAGATTGAAGAAATTAAGGAAAGAGCCAGCA 3495
Qy 173 ----- 173
Db 3496 CTCACAGAAACACAGCAACAACTCTTAATGTACAGAAAGAGATGAGTGACAGAA 3555
Qy 173 ----- 173
Db 3556 AAGATTAAATGAAATAGAGATTTAAAGATGAATTAAGAACAAAGAAATGACATTGGA 3615
Qy 174 -----AlaAlaLeuCysLysLysTyrAlaGluLeuLeuGluHis 187
Db 3616 CATATGGAACACAGAGAGCTTGAGTTGGCTCAGAACTTAATGAAATTTATGAGGAAGTG 3675
Qy 188 ArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGlu 207
Db 3676 AATCTATAACCAAGAAAGAAAGTTCTAAAGGAATTAAGAGTCAATTTGAAACACAG 3735
Qy 208 LysAspHisLeuArgGly----- 213
Db 3736 AGAGACCACCTTAGAGGATATATAAGAGAAATTTGAAGCTACAGGCTACAAACCAAGAA 3795
Qy 214 -----GluHisSerLysAlaVal-----LeuAla 221
Db 3796 GAACATAAAATGCTCATATTCACCTAAAGAAACACCAAGAACTATTGATGAACATAAG 3855
Qy 222 ArgSerLysLeuGluSerLeuCys-----ArgGluLeuGlnArgHisAsn 236
Db 3856 AGAGCGTATCTGAGAGACAGCTCAAAATAATAATTAATCTCAGGACTTGAAGAAATCCCAT 3915
Qy 237 ArgSerLeuLysGlnGluGlyValGlnArgAlaArgGluGluGluLysArgLysGlu 256
Db 3916 ACCAAATTAACAAGAGAG-----ATCCAGCTGCTTCATGAGGAACAGAGTTACTGCTAAT 3972
Qy 257 Val-----ThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGln 274
Db 3973 GTGAAAAAAGTCAGTGAGACTCAGAAACAAATGAATGAATGAGTGAATTAACAGACAG 4032
Qy 275 HisAsnGluArgAsnSer-----LysLeuArgGlnGluAsnMetGluLeuAla 290
Db 4033 TCCACAAACCAAGGACTCAACAAACACTGGCAGAGAAATGAATGAAGGCTCAGGTTGAAT 4092
Qy 291 GluArgLeuLysLysLeuIleGluGlnTyrGluLeuArgGluGluHisIleAspLysVal 310
Db 4093 GAAAAAATTCAGAAAGTCAGGAGAGAGATAAATCTCTAACCAAGAGAGACAACTT 4152

Qy 311 -----PheLysHisLysAspLeuGlnGlnLeuValAsp 322
Db 4153 AARACGTAARAAAGACCCCTTGAGCTTAAACATGACCAGCTGAAGAAACATATTAGAGAA 4212
Qy 323 -----AlaLysLeuGlnGlnAlaGln----- 329
Db 4213 ACTTTGGCTAAATCCAGAGTCTCAAGCAACAAAGAACAGTCCTTAAATATGAAGAA 4272
Qy 329 ----- 329
Db 4273 AAAGACAATGAACACTACCAAAATCGTGAAGTGAAGTGAAGTCAAAACCCCAAGATTCA 4332
Qy 330 -----GluMetLeu-----LysGluAlaGluGluArgHis 339
Db 4333 GCCTACTAAGGATAGAAATAGAAATGCTCGGATTTGCCAAAGAGCTTCAAGAAAGTCA 4392
Qy 340 GlnArgGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeu 359
Db 4393 GATGAAATGAATCT--GTAGCTAAGGAGAAAGATGACCTACAGAGGCTGCAAGAAGTT 4449
Qy 360 MetLysGlnGlnGluThrHisLysLysGlnGlnLeu-----AlaLeuTyrThr 375
Db 4450 CTTCAATCTCAAGAGTCACAGCTCAAGAGAAACATAAAAGAAATTTAGCTAAACCTG 4509
Qy 376 GluLysPheGluGluPheGln----- 382
Db 4510 GAACTCAAGAGAACTTAAAGTTGCTCATTTGCTCAAGAGAAACAGAGAACTATT 4569
Qy 383 -----AsnThrLeuSerLysSerSerGluValPheThrThrPhelLysGln 397
Db 4570 AATGAGTTAAGAGTGAATCTTTTCAAGAGAAAGAACTGAATATCAACCTTCAAGAGCAG 4629
Qy 398 -----GluMetGluLysMetThrLysLysIleLysLysLeu---GluLysGluThr 413
Db 4630 TTGAACCAATCAATGATAAATTACAGACAGATCCAGAGATTTATGAGAAAGGAA 4689
Qy 414 ThrMet-----TyrArgSerArgTyrGluSerSerAsnLysAlaLeuLeuGluMet 430
Db 4690 CAATTTAATATAAACAATTTAGTGAGTTTCAAGAAACGTTGAATGAACATTC 4749
Qy 431 AlaGluGluLysThrValArgAspLysGluLeuGluGlyLeuGlnValLysIle----- 448
Db 4750 AAGAGCATCGCAAGCCAGGATTCAGCCTCAACAGTATAGAAGTGAAGTGTGCGAG 4809
Qy 449 -----Gln 449
Db 4810 TTGACCAACAGACTTCAAGAAAGTCAAGAAAGAAATACAAATTTATGATTAAAGAAAAAGAG 4869
Qy 450 ArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgVal 469
Db 4870 GAATGMAAGAGTACAGAGGCCCTTTCAGATAGAGAGACCACTGAAGAAACACT 4929

RESULT 43

US-09-620-312D-130
; Sequence 130, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui

; APPLICANT: Zhou, Ping
 ; APPLICANT: Ma, Yundong
 ; APPLICANT: Wang, Dunrui
 ; APPLICANT: Wang, Zhiwei
 ; APPLICANT: John Tillinghast
 ; APPLICANT: Drmanac, Radoje T.
 ; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
 ; FILE REFERENCE: 784CIP2B
 ; CURRENT APPLICATION NUMBER: US/09/620,312D
 ; CURRENT FILING DATE: 2000-07-19
 ; PRIOR APPLICATION NUMBER: 09/552,317
 ; PRIOR FILING DATE: 2000-04-25
 ; PRIOR APPLICATION NUMBER: 09/488,725
 ; PRIOR FILING DATE: 2000-01-21
 ; NUMBER OF SEQ ID NOS: 1105
 ; SOFTWARE: pc_FL_Genes Version 1.0
 ; SEQ ID NO 130
 ; LENGTH: 8503
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (91)..(8082)
 ; US-09-620-312D-130

Alignment Scores:
 Pred. No.: 1-2e-10 Length: 8503
 Score: 239.00 Matches: 130
 Percent Similarity: 35.66% Conservative: 100
 Best Local Similarity: 20.16% Mismatches: 177
 Query Match: 8.85% Indels: 238
 DB: 4 Gaps: 25

US-10-023-523-8 (1-530) x US-09-620-312D-130 (1-8503)

QY	55	valserGluGluLeuSerArgGlnLeuGluAspIleLeuSerThrCysValAspAsn	74	3556	AAAGATTAAATAGAGAAATTTAAAGAAATGAATTAAAGAAACAAGAAATTCACATTGGAA	3615
Db	3049	ATTCTGAGGAAGATTTCAGGAATTTGCAT		174	-----AlaAlaLeuCysLysLysTyrAlaGluLeuLeuGluGluHis	187
QY	75	AsnGlnGlyProGlyGluAspGlyAlaGlnGly		3616	CATATGGAAACAGAGAGCGCTTGCTGCTCAGAACTTAATGAAATTTATGAGGAAGTG	3675
Db	3088	ATAACAGGAACAATAAGATGAATTCAGCAAAAGATGGTTGGCATAGATAAAACAG		188	ArgAsnSerGlnLysGlnMetLysLeuLeuGlnLysLysGlnSerGlnLeuValGlnGlu	207
QY	91	ProGluAspAlaGluLysSerArgThrTyrValAla		3676	AAATCTATAACCAAGAAGAAGAAAGTTCTAAGAAATTCACAGAGTCATTGGAACACAG	3735
Db	3148	GATTGGAGCTTAAATAATACCAACACTAATCGCATGATGTTAAGCAATAATGAGATAAT		208	LysAspHisLeuArgGly	213
QY	103	-----ArgAsnGlyGluProGluProThr		3736	AGAGACCACCTTAGAGGATATATAAGAGAAATTTGAAGCTACAGGCTTACAAACCAAGAA	3795
Db	3208	GAGCAACAAGGAAGATATTTCTTTAATACAGGAGAAATTAAGTCAACAAATGTTA		214	-----GluHisSerLysAlaVal	
QY	111	ProValValTyrGlyGluLysGluProSerLysGlyAsp		3796	GAACTAAAAATTCCTCATATTACCTAAAGAAACCAAGAACTATTGATGAACCTAAGA	3855
Db	3268	GAGAGTGTATAGACAGAAAGGAACAATTGAAGACTGACCTAAAGGAAATATTGAATG		222	ArgSerLysLeuGluSerLeuCys	
QY	124	-----ProAsnThrGluGluLeuArg		3856	AGAAGCGCTATCTGAGAAAGACAGCTCAAAATAATAATACTCAGGACTTAGAAAATCCCAT	3915
Db	3328	ACCATTGAAACCCAGGAAGAATTAAGACTTCTGGGATGAATTAATAAGACCAAGAG		237	ArgSerLeuLysGluGlyValGlnArgAlaArgGluGluGluLysArgLysGlu	256
QY	141	ArgArgProGlnGluLysLysAla		3916	ACCAATTACAGAAGAG--ATCCCGAGTCTTCATGAGAAACAAGAGTTACTGCCTAAT	3972
Db	3388	ATAGTTCACAAAGAAAGAACCATGCCATATAAGAAAGAGAGAG		257	Val-----ThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGluGln	274
QY	160	MetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeu		3973	GTGAAAAAGTTCAGTGAGACTCAGGAAACAATGATCACTGAGGTATTATTAACAGAACAG	4032
Db	3436	TCTAGGACCTGTGACAGACTGGCAGAGATTGAAGAAATTAAGAAAGAAAGAGCCAGCA		275	HisAsnGluArgAsnSer	290
QY	173	-----		4033	TCCACAACCAAGGACTCAACAACACTGGCAAGATAGAAATGCAAGGCTCAGGTTGAAT	4092
Db	3496	CTCCAGAAACACAGCAACACTTCTTAATGTACAGAGAGATGAGTGAGATGCAGAA		291	GluArgLeuLysLysLeuLeuGluGlnTyrGluLeuArgGluGluHisLysVal	310
QY	173	-----		4093	GAAAAATTTCAAGAAAGTCAGGAGAGATTAATCTCTAACCAAGGAAAGAGACAACCTT	4152
Db	3555	-----		311	-----PheLysHisLysAspLeuGlnGlnGlnLeuValAsp	322
QY	173	-----		4153	AAAAACGATAAAGAGCCCTTGAAGTTAAACATGACCACTGAAAGCAACATATTAGAGAA	4212
Db	3555	-----		323	-----AlaLysLeuGlnGlnAlaGln	329
QY	173	-----		4213	ACTTTGCTTAATATCCAGGAGTCTCAAGCAACAAGAAACAGTCCTTAAATATGAAGAA	4272
Db	3555	-----		329	-----	329
QY	173	-----		4273	AAAGACAATGAAACTACCAAAATCGTGAGTGAGATGAGCAATTCACAAACCAAGATTCA	4332
Db	3555	-----		330	-----GluMetLeu	339
QY	173	-----		4333	GCNCTACTAAGATAGAAATAGAAATGCTCGGATGTCGAAAGACTTCAAGAAAGTTCAT	4392
Db	3555	-----		340	GlnArgGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLeu	359
QY	173	-----		4393	GATGAAATGAAATCT--GTAGCTAAGGAGAAAGATGACCTACAGAGGCTCAAGAAAGTT	4449
Db	3555	-----		360	MetLysGlnGlnGluThrHisLeuLysGlnGlnLeu	375
QY	173	-----		4450	CITCAATCTGAAGATGACCAAGCTCAAGAAAAACATAAAAGAAATTTGTAGTAAACACCTG	4509
Db	3555	-----		376	GluLysPheGluGluPheGln	382
QY	173	-----		4510	GAACTGAAGGAACTTAAGTTCTCATTTGCTTGAAGAAACAGAGAAACTATT	4569
Db	3555	-----		383	-----AsnThrLeuSerLysSerGluValPheThrPheLysGln	397
QY	173	-----		4570	AATGATTAAGAGTGAATCTTTTCAGAGAGGAAACTGAAATATCAACCATTCAAAGACAG	4629
Db	3555	-----		398	-----GluMetGluLysMetThrLysLysLysLeu	413
QY	173	-----				

Db 4630 TTGAAGCAATCAATGATAAATATACAGAACACAGATCCAGAGATTTATGAGAAAGAGGAA 4689
 Qy 414 ThrMet-----TyrArgSerArgTrrpGluSerSerAsnLysAlaLeuLeuGluMet 430
 Db 4690 CAACCTTAATAAACAATAGTGGTTTCAGGAAAGAGTGAATGAATCAATCAATTC 4749
 Qy 431 AlaGluLysThrValArgAspLysGluLeuGluGluLysValLysle----- 448
 Db 4750 AAGAGCATCGCAAGCAAGGATTCAGCACTACAAAGTATAGAAAGTAAAGTCTCGAG 4809
 Qy 449 -----Gln 449
 Db 4810 TTGACCAACAGACTTCAAGAAAGTCCAGAGAAATACAAATTATGATTAAAGAAAGAG 4869
 Qy 450 ArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsnLysArgVal 469
 Db 4870 GAATGAAGAGTACAGGAGGCCCTTCAGATAGAGAGAGCAACTGAAGAAAGCACT 4929
 Qy 470 GlnAspLeuSerAla 474
 Db 4930 AAAGAAATTGAGCT 4944

RESULT 44

US-09-620-312D-1065
 ; Sequence 1065, Application US/09620312D
 ; Patent No. 6569662
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Asundi, Vinod
 ; APPLICANT: Zhang, Jie
 ; APPLICANT: Ren, Feiyan
 ; APPLICANT: Chen, Rui-hong
 ; APPLICANT: Zhao, Qing A.
 ; APPLICANT: Wehrman, Tom
 ; APPLICANT: Xue, Aidong J.
 ; APPLICANT: Yang, Yonghong
 ; APPLICANT: Zhou, Ping
 ; APPLICANT: Ma, Yunqing
 ; APPLICANT: Wang, Dunrui
 ; APPLICANT: Wang, Zhiwei
 ; APPLICANT: John Tillinghast
 ; APPLICANT: Drmanac, Radoje T.
 ; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
 ; FILE REFERENCE: 784CIP2B
 ; CURRENT APPLICATION NUMBER: US/09/620,312D
 ; PRIOR FILING DATE: 2000-07-19
 ; PRIOR FILING DATE: 2000-04-25
 ; PRIOR FILING DATE: 2000-01-21
 ; NUMBER OF SEQ ID NOS: 1105
 ; SOFTWARE: pt_FL_genes Version 1.0
 ; SEQ ID NO 1065
 ; LENGTH: 4892

TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (634)..(4275)
 US-09-620-312D-1065
 Alignment Scores:
 Pred. No.: 7.18e-11 Length: 4892
 Score: 237.50 Matches: 116
 Percent Similarity: 40.34% Conservative: 72
 Best Local Similarity: 24.89% Mismatches: 175
 Query Match: 8.79% Indels: 103
 DB: 4 Gaps: 17

US-10-023-523-8 (1-530) x US-09-620-312D-1065 (1-4892)

Qy 75 AsnGlnGlyGly-----ProGlyGluasp 82
 Db 646 AACCAAGTGTGAAAGTGCATCTGATGGCATATCTCTCTGCCAAGCCCTCCATCATC 705
 Qy 83 GlyAlaGlnGluProAlaGluProGluAspAlaGluLysSerArgThrTyrValAla 102
 Db 706 GGAATGCTCTGTAAGAAAGTCTCTCAGAGATGCAAAAGAAAGAAAG----- 753
 Qy 103 ArgAsnGlyGluProGluProThrProValValTyrGlyGluLysGluProSerLysGly 122
 Db 754 -----AAATCA 759
 Qy 123 AspProAsnThrGluGluLeuArgGlnSerAspGluValClyAspArgAspHisArg 142
 Db 760 AATAGGAGGAGGATGATGTCTATGGCTCAGGAACGTGTC-----AAACGACACCTAA 813
 Qy 143 -----ProGlnGluLysLysAlaLysGlyLeuGlyLysGlu----- 155
 Db 814 ACATCTGGAGAAATGCAACGAAACCTAAGAAATCCTGGAGTTATCCAAAGAACACCTC 873
 Qy 156 IleThrLeuLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAla 175
 Db 874 ATCCAACTACTCAGTATAATAGGAGGGAGTTCCAGGCCAGAGAGATGTGATCCACATG 933
 Qy 176 LeuCys-----LysLysTyrAlaGluLeuGluGluHisArgAsnSerGlnLysGln 193
 Db 934 CTGAAGACAGAGAAACCAAGCCTGAGGTTCTGGAGGCTCATTACGGGTCTCGGAGCCA 993
 Qy 194 MetLysLeuLeuGln-----LysLysGlnSerGlnLeuValGlnGluLysAspHis 211
 Db 994 GAGAAAGTGTCTGGGTCTGACCGAGATGCCATTCTTCCCGCAGAGAGAAATCCATAGA 1053
 Qy 212 ArgGlyGluHisSerLysAlaValLeuAlaArgSerLysLeuGluSerLeuGlu 231
 Db 1054 GAAGATGCTCTCAGAAACCGATTTCAGAGCTCGACAGACTTGAGGAAACACAGAAAGAA 1113
 Qy 232 LeuGlnArg-----HisAsnArgSerLeuLysGluGluGlyValGlnArgAla 248
 Db 1114 ACCTACCGCGCATCTAGAGCAGCTGTGTCGGCGAGAGTGTCTATAGGCGCACCGTA 1173
 Qy 249 GluGluGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAsp 268
 Db 1174 TAGAGTTAGAGAACGAGAGCATAAACACTGACTACTAGTACAGAACAGAGCGACCTC 1233
 Qy 269 GlnLeuGlnMetGluGlnHisGlnArgAsnSerLysLeuArgGlnGlu----- 285
 Db 1234 ACCAACCCTGCTGGAGGAGCGGAGAGGTTAAAAAGCTCCTTGAACAGAGAAAGGCT 1293
 Qy 286 -----AsnMetGluLeuAlaGluArgLysLysLeuLysLeuGluGln----- 299
 Db 1294 TATCAAGCCCGCAAGAAAGAAAGTAAACGACTCAATAAATAAGATGAGCTT 1353
 Qy 300 -----TyrGluLeuArgGluGluHisLeuLys 309
 Db 1354 GTTAAACTCAATCTTTCACACTCATGCTGGTGGATGAAGAGCAATGACATGACAT 1413
 Qy 310 ValPheLysHisLysAspLeuGlnGlnLeuValasp-----AlaLysLeuGln 326
 Db 1414 CTT-----GGCTGCAAGCCAGAAAGTACAGATCTTACTCAGAGCTGAGG 1461
 Qy 327 GlnAlaGlnGluMetLeuLysGluAlaGluArgHisGlnArgGluLysAspPheLeu 346
 Db 1462 GAAGAAGAGAGAGAGGCTCAAGCCATTACTTCCAAATCCAAAGAGAGAGAGAGAAATG 1521
 Qy 347 LeuLysGluAlaVal-----GluSerGlnArg 355
 Db 1522 CTCAGTTAGAACTGGACTTGAACACAGAGCTTCGAGTTTCTCAGAGCATGAGAG 1581
 Qy 356 MetCysGluLeuMetLysGlnGlnGlnThrHis-----LeuLysGlnGlnLeuAla 372
 Db 1582 ATGAACGCTAAACTGGCTAATCAAGAGTCTCAATAGGCAACTTAGCTCAAGCTGTT 1641

373 LeuTyrThrGluLysPheGluPheGlnAsnThr-----LeuSerLysSerSer 389
 1642 GGCTTACCCAGAAATCTCAGAGAGACCAACAAAAATCTCGAAGAGCGAG 1701
 390 GluValPheThrPheLysGlnGluMetGluLysMetThrLysLysLysLysLysLeu 409
 1702 GAA-----GMACTTCAGANTTAAGACATAAAATTGCCAAGGA 1740
 410 GluLysGluThrThrMetTyrArgSerArgTyrGluSerSerAsnLysLysAlaLeuLeuGlu 429
 1741 GAATGTGGAAACTCTAGCCCTCATCGGACAGAGTGGAAAAATCTCGAAAGCGTGTGCTTGA 1800
 430 MetAla-----GluGluLysThrValArgAspLysGluLeuGluLysLeuGlnVal 446
 1801 ATGGAAGGTAAGCATGAGGAGATCACTAAACCTGAATCCAGGTGAGGNAATTGCGAAG 1860
 447 LysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAspLeuAsn 466
 1861 AAGCTGCAGAGGAAGAACACCACTAGTAGGAGCTCAGACTTGAAGTTCGAAAGCTTACAG 1920
 467 LysArgValGlnAspLeu 472
 1921 AAGAGAAATGCTGAACTA 1938

RESULT 45

```

US-09-976-594-735
; Sequence 735, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 735
; LENGTH: 5467
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. 6673549 3151579CB1

```

Alignment Scores:			
Pred. No.:	9.36e-11	Length:	5467
Score:	237.001	Matches:	160
Percent Similarity:	38.71%	Conservative:	99
Percent Local Similarity:	23.92%	Mismatches:	217
Query Match:	8.77%	Indels:	194
DB:	4	Gaps:	34

US-10-023-523-8 (1-530) x US-09-976-594-735 (1-5467)

6	GlnProGlu	-----GlyProGluGlyAlaGlnGlu-----	16
2855	CAGGAAGAGGCTGGAGCTGAAGAGAGCTGGCGCAGAGCTGATCAAGTGTCACGG		2914
17	-----ArgProSerGlnAlaAlaProAlaValGluAlaGluGlyProGly	31	
2915	GGAGAAGCGGACTGCGCGAGCGGAGCTCTCGCGCTGCA-----CAGAGAGTGCCCGAGGT	2971	
32	SerSer-----GlnAlaProArgLysProGluGlyAlaGln-----	43	
2972	CCGTCAGCACAATGGCGGACCTTGAAGGCGATCTCCAGTCGCGCAGAAAGAGCGAGACGA	3031	
44	-----AlaArgThrAlaGlnSerGly-AlaLeuArgAs	54	
3032	GATGGAACAACACTTTCGAGTCGTCGATTCGATAAGGAGCAGATGTCGCGGTACAGA	3091	

Db 2209 GAGCATCCAGATGCGCCAGCAGCGCCAGCGCCAGTTCAGGACGACTTGGCGCAAGCCGA 2150
 Qy 291 uArgLeuLysLysLeuIleGluGlnTyr-----GluLe 302
 Db 2149 ACGGACCCGCGAGAAACCGAATCGACCTCGGGCATGACCGTACCTTGTCTCGGACCCCT 2090
 Qy 302 uArgGluGluHisIleAspLysValPheLysHisLysAspLeuGlnGlnLysValAs 322
 Db 2089 GGCCGAGAA-----ATGGCCATGCTCGCACCGGACAGCAAGCACTCAGCGC 2045
 Qy 322 pAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGlu-----GluArgHis 339
 Db 2044 GGCCGCCGCGAGAGAGCGGCATCGCCCTCGAACAGCGCGAGCGGCGATGACGCGCTG 1985
 Qy 339 sGlnArgGluLysAspPheLeuLeuLysGluAlaValGluSerGlnArgMetCysGluLe 359
 Db 1984 GCACGACGATGGATGCTTCAACACGACGAGGCGCCGAACCGCGCGCGCCGAGGT 1925
 Qy 359 uMetLysGlnGlnGlnThrHisLeuLysGlnGlnLeuAlaLeuTyrThrGluLysPheG 379
 Db 1924 GCACGATCGGCATCCAGCACCTCGGACGAGCGCTG-----1888
 Qy 379 uGluPheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMe 399
 Db 1888 -----1888
 Qy 399 tGluLysMetThrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 419
 Db 1887 -GAGCGCTGCAGATCGCGAGCGCGCTCGAGGAGCGGTGGCGAGTGGCGCC-- 1831
 Qy 419 gTrpGluSerSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLys 439
 Db 1830 ---GACCCGAGCAGCGCGCATCTCTGAACCTCAACAGACAGTGGCGATCGCGCAACT 1775
 Qy 439 sGluLeuGluGluGlnValLysLysLysLysLysLysLysLysLysLysLysLysLys 455
 Db 1774 GGCCCTGGAGAACTGCAATGTCAGGAGCAGGCGGCGGCAAGCTCGAACAAATTGG 1715
 Qy 455 sArgAlaLeuGlnThr-----GluArgAsnAspLe 465
 Db 1714 CCAGGAATTGCAGCAGCTGGCGCGCGCAACAGCACCGAGCGCGCGGCGGTGCGCGCCT 1655
 Qy 465 uAsnLysArgValGlnAspLeuSerAla-----GlyGlyGlnG 478
 Db 1654 GAACGGCGCATCGCTTCGCTGGAGGCGCTGCGACGAGCGCGCTCTGATCCCGGACAGG 1595
 Qy 478 ySer-----LeuThrAspSerGlyProGluArgProGlu-GlyProGlyAla- 494
 Db 1594 CGCCTTGGAGTGTTCGCGAGCAGCGGCTTGAACACAGCTCGCGCGCTCGCGAAGCCT 1535
 Qy 495 -----GlnAlaProSerSerProArgValThrGluA 505
 Db 1534 GCGTTCGAGCCGGTGGAGCTGGCGGTGGAGACCGTCTCGGCGCGGATCTGCAGGC 1475
 Qy 505 laProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGluP 525
 Db 1474 GCTTCTGCTG-----ACGGCTTCAGCGGCTCGCCCTGGC 1439
 Qy 525 roThrSerAlaArgAla 530
 Db 1438 CGGCTTCGCAAGGCG 1422

RESULT 47

US-09-252-991A-15688
 ; Sequence 15688, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 15688
 ; TYPE: DNA
 ; LENGTH: 3606
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-15688

Alignment Scores:
 Pred. No.: 7.6e-11 Length: 3606
 Score: 235.00 Matches: 144
 Percent Similarity: 39.87% Conservative: 106
 Best Local Similarity: 22.97% Mismatches: 219
 Query Match: 8.70% Indels: 158
 DB: 4 Gaps: 27

US-10-023-523-8 (1-530) x US-09-252-991A-15688 (1-3606)

Qy 3 SerProGlyGlnProGluAlaGlyProGluGlyAlaGlnGlu---ArgProSerGlnAla 21
 Db 120 GGCCTCGAAGACATCAAGCTGCG-----GGGTTCAAGTCTCTCGTGCATCGACAC 173
 Qy 22 AlaProAlaValGluAlaGluGlyProGlySerSerGlnAlaProArgLysProGluGly 41
 Db 174 GGTGAATCTCCCAACACATGCG---GGCAGTGTAGGGCCGAAACGCTCGGCAAGTC 230
 Qy 42 AlaGlnAlaArgThrAlaGlnSerGlyAlaLeuArgAspValSerGluGluLeuSerArg 61
 Db 231 GAACATCATCAACCGGTCGCTGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 290
 Qy 62 GlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsnGlnGlyGly---ProGly 80
 Db 291 CGAGTCGATGACCGATGTCATCTTCAACGGCTCGAATACCGCAAGCGGTGAGCCAGGC 350
 Qy 81 GluAspGlyAlaGln-GlyGluProAlaGlu-----ProGluAs 93
 Db 351 GAGCATCGAGCTGATCTTCGACACCGCGGACACACACTGGTGGCGGATAGCCCGAGTA 410
 Qy 93 pAlaGluLysSer---ArgThrTyrValAlaArgAsnGlyGluProGluProThrProVa 112
 Db 411 CGCCGATGATATTCATTCGCGCGGGTCTCGCGGATGGG-----CAGAACACTATT 464
 Qy 112 lValTyrGlyGluLysGluProSerLysGly-----122
 Db 465 CCTCAACCGCACCAAGTCCCGCGCGCGACATCACCGACATCTTCTCGGCACCGCGCT 524
 Qy 123 -AspProAsnThrGluGluIleArgGlnSerAspGluValGlyAspArgAspHisArg 142
 Db 525 GGGGCGCGCGAGTACTCGATCATCGAACAGGCGATGATCTCCAAGTGTATCGAGGCGCG 584
 Qy 142 gProGlnGluLysLys-----LysAlaLysGlyLeuGlyLysGluIleThrLe 158
 Db 585 TCCGGAAGACCTGCGCAACTTCATCGAGAAAGCGCGGCGCATTTCCAAGTACAAAGAGCG 644
 Qy 158 uLeuMetGlnThrLeuAsnThrLeuSerThrProGluGluLysLeuAlaLeuCys-- 177
 Db 645 CGCGCGCGAAACCGAAGCCGATCCGTCGCGACCCGAGAAACCTGGACCGCTCACCGA 704
 Qy 178 -----LysLysTyrAlaG 182
 Db 705 CCGTGGCGAAGAGTGGGGCGCAACTGGAACGCTGCACCGCGAGGCGGCGGCGGGA 764
 Qy 182 uLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLys-----LeuLe 197
 Db 765 AAAATACGAGGACAGAGCGCGGAGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 824
 Qy 197 uGlnLysLysGlnSerGlnLeuAlaGlnGluLysAspHisLeuArgGlyGluHisSerLy 217
 Db 825 CTGGCGCGACCTGAACGAGAGGAGGTGCGTCAAGCGGAGCGGCGGTGTCATCGCGCAGC 884

QY 217 sAlaValLeuAlaArgSerLysLeuGluSerLysLeuCysArgGluLeuGlnArgHisAsnAr 237
Db 885 CGCC-----TTCCAGGCGCTTGGTGGCGAG---CAGCGGGT----- 918
QY 237 gSerLeuLysGluGluGlyValGlnArgAlaArg-----GluGluGluGluGly 253
Db 919 -----GCCATCCCGCATCGCAACGCTCGCGAGCGGCATCAGCATGTCGCAAGC 971
QY 253 sArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeu-----G1 271
Db 972 CTTCAACAGGTGCAGGACGCTTCTATTTCGTCGGCGGCGACATCGCCGGGTGCGAGCA 1031
QY 271 nMetGluGlnHisAsnGluArgAsnSerLysLeuArgGlnGlnAsnMetGluLeuAlaG 291
Db 1032 GAGCATCAGCATCGCCAGCGCCGCGCATGTCAGGACGACTTGCAGGAGCCGA 1091
QY 291 uArgLeuLysLysLeuIleGluGlnThr-----GluLeu 302
Db 1092 ACGGACCGCCAGGAAACCGAATCGCACTCGGCGATGACGCTTGTCTCGGACCT 1151
QY 302 uArgGluGlnHisIleAspLysValPheLysHisLysAspLeuGlnGlnLeuValAs 322
Db 1152 GGCGGAGAA-----ATGGCCATGCTCGCACCGGACGAGCACTCAGCGC 1196
QY 322 pAlaLysLeuGlnGlnAlaGlnGluMetLeuLysGluAlaGlu-----GluArgHi 339
Db 1197 GGCCGCGCGCAAGCGGCGATCGCTTGGACAGCGCGGAGCGGATGCGAGCGCTG 1256
QY 339 sGlnArgGluLysAspPheLeuLysGluAlaValGluSerGlnArgMetCysGluLeu 359
Db 1257 GCAGCAGCGTGGGATGCTTCAACAGCAGAGCGCGGACCCCGCGCGGCGGAGGT 1316
QY 359 uMetLysGlnGlnGlnThrHisLysGlnGlnLeuAlaLeuThrThrGluLysPheG1 379
Db 1317 GCAGCAGTGGCGATCCAGCACCTGGAGAGAGCGCTG----- 1353
QY 379 uGluPheGlnAsnThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluWe 399
Db 1353 ----- 1353
QY 399 tGluLysMetThrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 419
Db 1354 -GAGCGCTGCAGGATCGCGGCGGCGCTTCAGGAGGCGGCGGCGGCGGCGGCGGCGG 1410
QY 419 gTrpGluSerSerAsnLysAlaLeuLeuGluMetAlaGluLysThrValArgAspLys 439
Db 1411 ---GACCCCGAGCAGCGCGGATCTCTCGAACTCAACGAGCAGCGGCGGCGGCGGCGG 1466
QY 439 sGluLeuGluGlyLeuGlnValLysLysLysLysLysLysLysLysLysLysLysLys 455
Db 1467 GGCCCTGGAAGACTGCAATTGAGGAGCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1526
QY 455 sArgAlaLeuGlnThr-----GluArgAsnAspLys 465
Db 1527 CCAGGAATTGCAGCAGCTGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1586
QY 465 uAsnLysArgValGlnAspLeuSerAla-----GlyGlyGlnG1 478
Db 1587 GAACGGCGCATCGCTTCGCTGGAGGCGCTTCGAGCGGCGGCGGCGGCGGCGGCGGCGG 1646
QY 478 ySer-----LeuThrAspSerGlyProGluArgArgProGlu-GlyProGlyAla- 494
Db 1647 CGCCTTGGAGTGGTTCGCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1706
QY 495 -----GlnAlaProSerSerProArgValThrGluA 505
Db 1707 GCGTGTGAGCCCGGTTGGAGCTGGCGGTGGAGACGCTGCTCGGCGGCGGCGGCGGCGG 1766
QY 505 laProCysTyrProGlyAlaProSerThrGluAlaSerGlyGlnThrGlyProGlnGluP 525
Db 1767 GGTCTTCTGCG-----ACGGCTTTCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1802
QY 525 roThrSerAlaArgAla 530

Db 1803 CGGCTTCGCAAGGGCG 1819
RESULT 48
US-08-875-435B-1
; Sequence 1, Application US/08875435B
; Patent No. 6593304
; GENERAL INFORMATION:
; APPLICANT: Hasegawa, Kazuhide
; APPLICANT: Arakawa, Emi
; APPLICANT: Oda, Shoji
; APPLICANT: Matsuda, Yuzuru
; APPLICANT: Takahashi, Katsuhito
; APPLICANT: Sugahara, Michihiro
; APPLICANT: Ishiyama, Haruo
; TITLE OF INVENTION: RECOMBINANT DNA COMPRISING DNA CODING
; TITLE OF INVENTION: FOR MYOSIN HEAVY CHAIN SMI ISOFORM PROTEIN INSERTED INTO
; TITLE OF INVENTION: VECTOR DNA, MICROORGANISM CARRYING THE RECOMBINANT DNA, AND
; TITLE OF INVENTION: AN AGENT FOR TREATMENT OF ARTERIOSCLEROSIS COMPRISING THE
; TITLE OF INVENTION: RECOMBINANT DNA
; FILE REFERENCE: 07898-013001
; CURRENT APPLICATION NUMBER: US/08/875,435B
; CURRENT FILING DATE: 1997-07-25
; PRIOR APPLICATION NUMBER: PCT/JP96/00134
; PRIOR FILING DATE: 1996-01-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 6175
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (105)....(6020)
US-08-875-435B-1
Alignment Scores:
Pred. No.: 2,28e-10 Length: 6175
Score: 233.50 Matches: 154
Percent Similarity: 38.13% Conservative: 95
Best Local Similarity: 23.58% Mismatches: 237
Query Match: 8.64% Indels: 169
DB: 4 Gaps: 29
US-10-023-523-8 (1-530) x US-08-875-435B-1 (1-6175)
QY 5 GlyGlnProGluAlaGlyProGluGlyAlaGlnGluArgProSerGlnAlaProAla 24
Db 4058 GGACACCAAGAGCTGCTCCA-----AGAGAAACCCGCGAGAGCTCAATGTGC 4108
QY 25 ValGluAlaGluGlyProGlySer---SerGlnAlaProArgLysProGluGlyAlaGln 43
Db 4109 TACCAAGCTCGCTCAGTTGGAGATGAAGAGAACAGCCCTGCGAGGACCTGAT--GAG 4166
QY 44 AlaArgThrAlaGlnSerGlyAlaLeuArgAspValSer-----Glu 57
Db 4167 GAGATGAGGCTTAAGCAAAACCTGGAGCGCCATGTCTCAACTGAACATTCAGCTCTCA 4226
QY 58 GluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnGlnGly 77
Db 4227 GACTCTAAGAGAGCTGCGAGGACTTTGCAAGTACCATCGAGTCTC----- 4271
QY 78 GlyProGlyLysGlyAlaGlnGlyGluProAlaGluProGluAspAlaGlyLysSer 97
Db 4272 -----ATGCGAGCGGGAAGAGAGGTTACAGAAAGAGATGAGGGCGCTC-----AGC 4319
QY 98 ArgThrTyrValAlaArgAsnGlyGluProGluProThrProValValTyrGlyGluLys 117
Db 4320 CAGCAGTAT-----GAGGAGAGCGCGGCTGCCCTATGACAACTG 4358
QY 118 GluProSerLysGlyAspProAsnThrGluGluIleArgGlnSerAspGluValGlyAsp 137
Db 4359 GAGAAACCAAG-----AACAGGCTCCAGCAGGAGCTGGATGACCTGGTGTGGAC 4409

Db 1011 GAAGCAACTCGAGCAGGAATCTCGGCACCTCGAGCAGCAGCAGCGGCGCAGCGCA 1070
Qy 56 SerGluGluLeuSerArgGlnLeuGluAspIleLeuSerThrTyrCysValAspAsnAsn 75
Db 1071 GCAAGCGCGCGCGCGCGCGCTTGCAAGATTGCGAGGCCGCTTCGAGGCGCGCA 1130
Qy 76 GlnGlyGlyProGlyGluAspGly-----AlaGlnGlyGlu 87
Db 1131 CGCGCAGCG 1190
Qy 88 ProAla-GluProGluAspAlaGluLysSerArgThrTyrValAlaAsnGlyGluPr 107
Db 1191 CCAGCTGGAGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1249
Qy 107 o-----GluProThrPro-----ValValTyrGlu 115
Db 1250 ATCGGCAATGAACCG 1309
Qy 115 yGluLysGluProSerLysGly-----AspProAs 125
Db 1310 AGCAGCAGCG 1369
Qy 125 nThrGluGluLeuArgGln-----SerAspGluValGlyAspArgAspHisArgPr 143
Db 1370 CTCTCG 1428
Qy 143 oGlnGluLysLysLysAlaLysGlyLeuGlyLysGluLeuThrLeuLeuMetGlnThrLe 163
Db 1429 -CAAGCTGCGCAGCG 1487
Qy 163 uAsnThrLeuSerThrProGluGluLysLeuAlaA----- 175
Db 1488 GGCCCG 1547
Qy 176 -----LeuCysLysLysTyrAlaGluLeuGluGluLeuGluGluGluGluGlu 191
Db 1548 CGAGCGCGCAGCAGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1607
Qy 191 nLysGlnMetLysLeuGlnLysLysGlnSerGlnLeuValGln----- 206
Db 1608 ACGCAACAGCGCTTCACAGGAATGGCAACAGCGCGCGCGCGCGCGCGCGCGCGCG 1667
Qy 207 -----GluLysAspHisLeuArgGlyGluHisLysAlaValLeuAlaAr 222
Db 1668 TCGAGCGCTGGCG 1727
Qy 222 gSerLysLeuGluSerLysCys----- 229
Db 1728 GAAACGACAGCG 1787
Qy 230 ---ArgGluLeuGlnArgHisAsnArgSerLeuLysGluGlyValGlnArgAlaAr 248
Db 1788 CACCGCGAGTGTCTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1847
Qy 248 g----- 248
Db 1848 CGCGCGCGCTGTGAGCG 1907
Qy 249 -----GluGluGluLysAr 254
Db 1908 CCATTCGAGCACTGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1967
Qy 254 gLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuGlnMetGlu 274
Db 1968 CGAGCAGTCCCTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2027
Qy 274 nHisAsnGluArgAsnSerLysLeuArgGlnGlnAsnMetGluLeuAlaGluArgLeu 294
Db 2028 CCACGGGAAAGATCTCAGAGCTGCCAGGAGCGAGCAGCACTGCTGCGCGCGCGCG 2087
Qy 294 sLysLeu-----IleGluGlnTyrGluLeuArgGluGluHisIleAspLysValPh 311

Db 2088 CGCGCTCATCGCACTGCACTGCACTGCACTGCGCGCGCGCGCGCGCGCGCGCG 2130
Qy 311 eLysHisLysAspLeuGlnGlnLeuValAspAlaLysLeuGlnGlnAlaGlnGluMe 331
Db 2131 -----CTGCACTCTCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2177
Qy 331 tLeuLysGlu-----AlaGluGluArgHisGlnArgGluLysAspPhe-----LeuLeuLys 348
Db 2178 CTTGAGCAGCTTGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2237
Qy 348 sGluAlaValGlnSerGlnArgMetCysGluLeuMetLysGlnGlnGluThrHisLeuLys 368
Db 2238 CGGCCAACCGCAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2297
Qy 368 sGlnGlnLeuAlaLeuTyrThrGluLysPheGluGluPheGlnAsnThrLeuSerLys 388
Db 2298 ACAGCGCAGCAGCG 2343
Qy 388 rSerGluValPheThrThrPheLysGlnGluMetGluLysMetThrLysLysLeuLys 408
Db 2344 -----TTGCAAGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2384
Qy 408 sLeuGluLysGluThrThrMetTyrArgSerArgTyrGluSer-----SerAs 424
Db 2385 ATTGCTCTCCGACGAG-----CACTGGCAGCGCGCGCGCGCGCGCGCGCGCG 2438
Qy 424 nLysAlaLeuLeuGluMetAlaGluGluLysThrValArgAspLysGluLeuGluLys 444
Db 2439 CCAGCAGCTCGCGCGAGTCCATCGAGCAACCGCGCGCGCGCGCGCGCGCGCGCG 2498
Qy 444 uGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluArgAsnAs 464
Db 2499 CGAACCGCGCTCTGAGGAGCTCAAGCAGCGCGCGCGCGCGCGCGCGCGCGCG 2554
Qy 464 pLeuAsnLysArgValGlnAspLeuSerAlaGlyGlyGlnGlySer----- 479
Db 2555 AGCAACCGCAGCAGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2614
Qy 480 -----LeuThrAspSerGlyProGluArgArgProGluGlyProGlyAlaGlnAlaPro 497
Db 2615 TCGCGCACTGAACG 2671
Qy 498 -----SerSerProArgValThrGluAlaProCys-----TyrPr 509
Db 2672 ATTGGCACTATCGCTGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2731
Qy 509 oGlyAlaProSerThrGluAlaSerGlyGlnThr-----GlyProGlnGluPr 525
Db 2732 AGCGCGCGCTGGATTCTGCGCGGAGCAACTGCGACTCGCGGAGCGCGCGCGCG 2791
Qy 525 oThrSerAlaArgAla 530
Db 2792 TGCAGCAGCAACGCGCA 2807

RESULT 50

US-09-010-147B-17
; Sequence 17, Application US/09010147B
; Patent No. 6653445
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Human Proteins
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC
; compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION RELEASE: US/09/010.147B

APPLICATION NUMBER: 6553445-2002

FILING DATE: 12-JAN-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/034,205

FILING DATE: 21-JAN-1997

APPLICATION NUMBER: US 60/034,204

FILING DATE: 21-JAN-1997

ATTORNEY/AGENT INFORMATION:

NAME: Jonathan L. Klein

REGISTRATION NUMBER: 41,119

REFERENCE/DOCKET NUMBER: PF353

TELECOMMUNICATION INFORMATION:

TELEPHONE: 301-309-8504

TELEFAX: 301-309-8439

INFORMATION FOR SEQ ID NO: 17:

SEQUENCE CHARACTERISTICS:

LENGTH: 2934 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 1..2934

SEQUENCE DESCRIPTION: SEQ ID NO: 17:

US-09-010-147B-17

Alignment Scores:

Pred. No.: 1.4e-10 Length: 2934
Score: 230.50 Matches: 136
Percent Similarity: 39.45% Conservative: 109
Best Local Similarity: 21.90% Mismatches: 173
Query Match: 8.53% Indels: 204
DB: 4 Gaps: 28

US-10-023-523-8 (1-530) x US-09-010-147B-17 (1-2934)

QY 31 GlySerSerGlnAlaProArgLysProGluGlyValAlaGlnAla 48
DB 786 GCGCATTCGAAAGCCAGCTGGAGAGAGAGAACTGCTGCCACACAGAGAGA 845
QY 49 SerGlyAlaLeuArgAspValSerGlu-GluLeuSerArgGlnLeuGluAspIleLeuSe 68
DB 846 TGGCGTGTGGCCAGAGCAAACTGAGGAGCTCAACAAGGAGATG- 891
QY 68 rThrTyrcysValAspAsnGlnGlyGlyProGlyGluAspGlyAlaGlnGlyLysPr 88
DB 892 -----GCAGCAAAAGGCCAA 908
QY 88 oAlaGluProGluAspAlaGluLysSerArgThrTyrrValAlaAa:gaSnglyGluProG1 108
DB 909 AGCAGAGCGCGGGAGGCCAAAGTCAAAAGAGAGCTGTGGTGGCCCG- 962
QY 108 uProThrProVal-----ValTyrrGlyGlu----- 116
DB 963 GATCAGCGTGTGCAGGACGATGACGAGCCAGTACCGGAGACGATGAAGGAGGTGCA 1022
QY 117 -----LysGluProSerLysGlyAspProAsnThrG1 127
DB 1023 GCAGCTGCAGGCAAGATCCGAGCTCTTCAGGAGCAGCTGGAGATGCCCCAACACGCA 1082
QY 127 uGluIleArgGlnSerAspGluValGlyAspArgAspHisArgArgProGlnGluLysLy 147
DB 1083 GTGGCCCCC----- 1092
QY 147 sLysAlaLysGlyLeuGlyLysGluIleThrLeuLeuMetGlnThrLeuAsn----- 164
DB 1093 -----CTGCAGAGAGAGAACTCCATCTCTGCGGGATGCTTGAACAGGCCAC 1139

QY 165 -----ThrLeuSerThrProGluGluLysLeuAlaA1 175
DB 1140 GAGCCAGGTGGAGAGCAAGCAAGAGAGCTGCCCAAGCTTCGCGAGAGCTCAGCAA 1199
QY 175 aLeuCysLysLysTyrrAlaGluLeuLeuGluGluHisArgAsnSerGlnLysGlnMetLy 195
DB 1200 GGTCAAGAAAGAGCTGGTGGAGAGTCTGGCGCAAGATGAGCAGCAGCGGAA 1259
QY 195 sLeuLeuGlnLysLysGlnSerGlnLeuValGlnGluLysAspHisLeuArgGlyGluHi 215
DB 1260 AGCTCTGAAAGCCAGCAGCTGCTTCGAGAGCAGGTCTCGAGCGGAGGTGCCA 1319
QY 215 sSerLysAlaValLeuAla---ArgSerLysLeuGluSerLeuCysArgGluLeu----- 232
DB 1320 CAGGAGAGTGGAGGAGGCTTCAGAGCGCTGGAGAGTGGAGCGGAGCTGTGCCA 1379
QY 233 ----GlnArgHisAsnArgSerLeuLysGluGluGlyValGlnArgAlaArgGluGluG1 251
DB 1380 CAGCAGAGCAGCAGCCAGCTCCGCGCGGAT---GCCAGAGAGGCCAGGAGCAACA 1436
QY 251 uGluLysArgLysGluValThrSerHisPheGlnValThrLeuAsnAspIleGlnLeuG1 271
DB 1437 GCAGCAGATGCCAGCTGCACACAAATTCACATCTCCGAGCGGAGGTGCCACAA 1496
QY 271 nMetGluGlnHisAsnGluArgAsnSerLysLeu-----ArgGlnGluAsnMetG1 288
DB 1497 ATGGAGAGCTGAGTGGCTCCACGCGCAGCTCCAGGAGGCCAGGCGAGAACTCCCA 1556
QY 288 uLeuAlaGluArgLeuLysLys-----LeuIleGluGlnTyrrGluLeuArgGluG1 305
DB 1557 GCTCAGAGAGAAATCCGTTCCATTGAGGCGCTCTGGAGCGGCGGCGCGGGATGC 1616
QY 305 uHisIleAspLysValPheLysHisAspLeuGlnGlnGlnLeuValAspAlaLysLe 325
DB 1617 C-----CAGGAGCTCCAGGCCAGCCAGCGGCGGAGCTGACCA 1652
QY 325 uGlnGlnAlaGlnGluMetLeuLysGluAlaGluGluArg---HisGlnArgGluLysAs 344
DB 1653 GCACGAGACTGCG-----CTCAGGAGCTGGAGTCCCGGTCTCGGAGTCTGGAGAAGA 1706
QY 344 pPheLeu---LeuLysGluAlaValGluSerGlnArgMet----- 356
DB 1707 GGCCATCAGCTCAGGAGGCGCTTCGAGCAGCAGAGAAAGTGAAGAACATACCTCCGGA 1766
QY 357 -----CysGluLeuMetLy 361
DB 1767 GAAGAACTGAAGGCCATGGAGCAGCTGGCCACCGCGGAGCAGGCTGCAAGGAGAAGCT 1826
QY 361 sGlnGlnGluThrHisLeuLysGlnGln----- 370
DB 1827 GCATCTCCTCAGCCAGCCAGGAGGAATCGGAGAGAGAGCTCTGTCTGATTGAGCGCA 1886
QY 371 -----LeuAlaLeu-----Ty 374
DB 1887 GACCATGAGGCGCTGTGCTCTGCTCCAGAACTCTGTCTTGGCACAACAGAAATTA 1946
QY 374 rThrGlu----- 376
DB 1947 CACCGATGCTGCAGGATCTCAAAGAGAAAGGCCACCGCTGCTGAAGCACCCGCGCAGC 2006
QY 377 -----LysPheGluGluPheGlnAs 383
DB 2007 TCCCGGAGGCCCTCTCGAGCTGCGCTCCAAAGTTGAGGAGGCGGAGGAGCAGCAGAG 2066
QY 383 nThrLeuSerLysSerSerGluValPheThrThrPheLysGlnGluMetGluLysMetTh 403
DB 2067 CACACTCAGGCGCAGTGTGACAGTACCAGCATCTCTGGCGGAGCAGCAGGAGCATGTCT 2126
QY 403 rLysLysIleLysLys---LeuGluLysGluThrThrMetTyrrArgSerArgTrpGluSe 422
DB 2127 CAGAGACCTGCAGAGAGCGTGGAGAGGAGGAGCAGAGTGTGGAGGGCCAG----- 2178
QY 422 rSerAsnLysAlaLeuLeuGluMetAlaGluGluLysThrValAa:gsAspLysGluLeuG1 442

```
Db 2179 -----GTGGGCGCGCAGAGAG-----GAGCTCCA 2204
QY 442 uGlyLeuGlnValLysIleGlnArgLeuGluLysLeuCysArgAlaLeuGlnThrGluAr 462
Db 2205 GAAGTCCCGGTTCACAGTCAGCATCTCGAGAGATTGTAGAGAAAGCTAAAGAGGAGAACT 2264
QY 462 gAsnAspLeuAsnLys--ArgValGlnAspLeuSerAlaGlyGlyGlnGlySerLeuTh 481
Db 2265 TGAAGTTCGGACCGAGGTGAGGA-GCACACGTTCGATTTTGGAGGCGAGCTGGGAAAGC 2323
QY 481 rAspSerGlyProGluArgArg---ProGlu-----GlyProGlyAlaGlnAlaPr 497
Db 2324 ACATGCGCGCGCCACGCGGAGTGCCAGAACTACGCCCAAGGAGGTGGCGGGCTGAGGC 2383
QY 497 oSerSerProArgValThr---GluAlaProCys---TyrProGlyAlaProSerThrCl 515
Db 2384 AACTTCTCTAGAAATCTCAATCTCAGCTCGATCGCGCCCAAGAGCGGAAAGCCAGAAACAGA 2443
QY 515 u 515
Db 2444 G 2444
```

Search completed: June 8, 2004, 12:41:47
Job time : 314.29 secs

